

APPLICANT: Leong, Kahan
APPLICANT: Raitano, Arthur B.
APPLICANT: Saffran, Douglas C.
APPLICANT: Yakobovits, Aya
TITLE OF INVENTION: BPC-1: A SECRETED BRAIN-SPECIFIC PROTEIN EXPRESSED AND
FILE REFERENCE: 1703-017.US1
CURRENT APPLICATION NUMBER: US/09/374,135
PRIORITY FILING DATE: 1999-08-10
PRIORITY FILING DATE: 1998-08-10
PRIORITY FILING DATE: 1998-08-10
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 3
LENGTH: 115
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-374-135-3

Query Match 30.2%; Score 260; DB 3; Length 115;
Best Local Similarity 43.5%; Pred. No. 1,1e-24;
Matches 50; Conservative 20; Mismatches 37; Indels 8; Gaps 2;

QY 52 IFTSPNYSKYPPDRECIYIIAARQCIELYFDEKYSIEPSW-----ECKFDHIEVR 104
DB 1 IFTSPNYSKYPPDRECIYIIAARQCIELYFDEKYSIEPSW-----ECKFDHIEVR 104
DB 1 IFTSPNYSKYPPDRECIYIIAARQCIELYFDEKYSIEPSW-----ECKFDHIEVR 104
QY 105 DGRFGSPILGRFCGQONPP-VIKSSGRFLMIKFFAGDELSMGFSARVNTPRK 158
DB 61 DGRFGSPILGRFCGQONPP-VIKSSGRFLMIKFFAGDELSMGFSARVNTPRK 158
DB 61 DGRFGSPILGRFCGQONPP-VIKSSGRFLMIKFFAGDELSMGFSARVNTPRK 158

RESULT 3

US-08-866-650-3
Sequence 3, Application US/08866650
Patent No. 5939321
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
APPLICANT: Takahara, Kazuhiko
APPLICANT: Hoffman, Guy G
TITLE OF INVENTION: Mammalian Tolloid-Like Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,650
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296, 93839
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-650-3

Query Match 27.7%; Score 238.5; DB 2; Length 1013;
Best Local Similarity 39.3%; Pred. No. 1.2e-20;
Matches 42; Conservative 25; Mismatches 39; Indels 1; Gaps 1;

QY 47 HAEGIFTSPNYSKYPPDRECIYIIAARQCIELYFDEKYSIEPSWCKFDHIEVRD 106
DB 779 HSPSGILTSFNPWDKYSRREKCTWISAIIGHRTTLAFNE-FVEHQEQCAVDHLEIFDG 837
QY 107 PGRFSPILGRFCGQONPP-VIKSSGRFLMIKFFAGDELSMGFSARVNTPRK 153
DB 838 ETEKSPILGRFCGQONPP-VIKSSGRFLMIKFFAGDELSMGFSARVNTPRK 153
DB 838 ETEKSPILGRFCGQONPP-VIKSSGRFLMIKFFAGDELSMGFSARVNTPRK 153

RESULT 4

US-09-021-287-3
Sequence 3, Application US/09021287
Patent No. 5981717
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
APPLICANT: Takahara, Kazuhiko
APPLICANT: Hoffman, Guy G
TITLE OF INVENTION: Mammalian Tolloid-Like Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,287
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/866,650
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296, 93839
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-021-287-3

Query Match 27.7%; Score 238.5; DB 2; Length 1013;
Best Local Similarity 39.3%; Pred. No. 1.2e-20;
Matches 42; Conservative 25; Mismatches 39; Indels 1; Gaps 1;

QY 47 HAEGIFTSPNYSKYPPDRECIYIIAARQCIELYFDEKYSIEPSWCKFDHIEVRD 106
DB 779 HSPSGILTSFNPWDKYSRREKCTWISAIIGHRTTLAFNE-FVEHQEQCAVDHLEIFDG 837
QY 107 PGRFSPILGRFCGQONPP-VIKSSGRFLMIKFFAGDELSMGFSARVNTPRK 153
DB 838 ETEKSPILGRFCGQONPP-VIKSSGRFLMIKFFAGDELSMGFSARVNTPRK 153
DB 838 ETEKSPILGRFCGQONPP-VIKSSGRFLMIKFFAGDELSMGFSARVNTPRK 153

RESULT 5
US-09-240-473-3
Sequence 3, Application US/09240473

```
; Patent No. 6297011
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hoffman, Guy G
; TITLE OF INVENTION: Mammalian Tollold-like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/240,473
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1013 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-240-473-3

Query Match      27.7%; Score 238.5; DB 3; Length 1013;
Best Local Similarity 39.3%; Pred. No. 1.2e-20;
Matches 42; Conservative 25; Mismatches 39; Indels 1; Gaps 1;

QY 47 HAEGIFTSNPVPSKYPPDRRCIYITBAAPQCITLYPDEKXSTEPSMECKPDHIEVDDG 106
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 779 HSPGSLITSPWPDYPRKRCITWISATPGHRIKLAPE-PEVQHOBCAYDHLIEFDG 837
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 107 PFGRSPITGRFCGOQNPPIKSSGRFLWIKFFADGELSMGFSARYN 153
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 838 ETEKSPILGRLOGSKIPDPLMATGNEMFIRFISDASVQRKGFQATHS 884

RESULT 6
US-08-866-650-5
; Sequence 5, Application US/0886650
; Patent No. 5939321
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hoffman, Guy G
; TITLE OF INVENTION: Mammalian Tollold-like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,650
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1013 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-866-650-5

Query Match      27.1%; Score 233.5; DB 2; Length 1013;
Best Local Similarity 40.2%; Pred. No. 4.9e-20;
Matches 43; Conservative 23; Mismatches 40; Indels 1; Gaps 1;

QY 47 HAEGIFTSNPVPSKYPPDRRCIYITBAAPQCITLYPDEKXSTEPSMECKPDHIEVDDG 106
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 779 HSPGSLITSPWPDYPRKRCITWISATPGHRIKLAPE-PEVQHOBCAYDHLIEFDG 837
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 107 PFGRSPITGRFCGOQNPPIKSSGRFLWIKFFADGELSMGFSARYN 153
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 838 ETEKSPILGRLOGSKIPDPLVATGNKMFVRVSDASVQRKGFQATHS 884

RESULT 7
US-09-021-287-5
; Sequence 5, Application US/09021287
; Patent No. 5981717
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hoffman, Guy G
; TITLE OF INVENTION: Mammalian Tollold-like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,287
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/866,650
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1013 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on:

May 18, 2004, 18:15:56 ; Search time 57 Seconds
(without alignments)
783.202 Million cell updates/sec

Title:
US-09-887-593a-2

Perfect score:
1 MHRKSVLHIVASLIILHLS.....ADGELSMGFSARNYTPGK 158

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:
1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: Genesegp1980s:*
2: Genesegp1990s:*
3: Genesegp2000s:*
4: Genesegp2001s:*
5: Genesegp2002s:*
6: Genesegp2003as:*
7: Genesegp2003bs:*
8: Genesegp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	862	100.0	158	3	AAy82894 Human bra
2	851	98.7	526	6	ABU69144 Human NOV
3	851	98.7	533	5	ABU79587 Human CUB
4	802	93.0	525	7	ADc42850 REMAP pro
5	479	55.6	130	4	ABBI10345 Human CDN
6	479	55.6	130	5	ABP66932 Human pol
7	479	55.6	141	4	AAU19919 Human hum
8	470	54.5	352	4	ABBI1868 Human sec
9	470	54.5	487	7	ADc42852 REMAP pro
10	470	54.5	525	4	AAU12228 Human pol
11	470	54.5	525	4	AAU12228 Human pol
12	470	54.5	525	4	AAU12228 Human pol
13	470	54.5	525	4	AAU12228 Human pol
14	470	54.5	525	6	ABU80926 Human PRO
15	470	54.5	525	6	ABU66626 Human PRO
16	470	54.5	525	6	ABU59707 Novel sec
17	470	54.5	525	6	ABO24897 Human sec
18	470	54.5	525	6	ABU66902 Human sec
19	470	54.5	525	6	ADA45633 Novel hum
20	470	54.5	525	6	ADA26938 Human PRO
21	470	54.5	525	6	ADA76064 Human PRO
22	470	54.5	525	6	ADA18714 Human PRO
23	470	54.5	525	6	ADA61337 Homo sapi
24	470	54.5	525	6	ADBI1922 Novel hum
25	470	54.5	525	6	ADBI27663 Human PRO

26	470	54.5	525	6	ADBI86142 Novel hum
27	470	54.5	525	6	ADBI15706 Human PRO
28	470	54.5	525	6	ADA47492 Human PRO
29	470	54.5	525	6	ADA67287 Human PRO
30	470	54.5	525	6	ADBI30294 Human PRO
31	470	54.5	525	6	ADBI85590 Novel hum
32	470	54.5	525	6	ADA96802 Human PRO
33	470	54.5	525	6	ADA79106 Human PRO
34	470	54.5	525	6	ADA87245 Novel hum
35	470	54.5	525	6	ADBI16447 Human PRO
36	470	54.5	525	6	ADA91539 Novel hum
37	470	54.5	525	6	ADBI14602 Human PRO
38	470	54.5	525	6	ADBI18563 Human PRO
39	470	54.5	525	6	ADA93778 Human PRO
40	470	54.5	525	6	ADBI19674 Novel hum
41	470	54.5	525	6	ADBI12986 Human PRO
42	470	54.5	525	6	ABO43205 Novel hum
43	470	54.5	525	6	ADBI74240 Human PRO
44	470	54.5	525	6	ADBI24473 Human PRO
45	470	54.5	525	6	ADBI1997 Human PRO

ALIGNMENTS

RESULT 1
ID AAY82894 standard; protein; 158 AA.
AC AAY82894;
DT 25-JUL-2000 (first entry)
DE Human brain specific protein BPC-1.
XX BPC-1; oncogene; oncogenic; cancer; prostate; bladder; antibody;
XX antisense; vaccine; detection; prognosis; drug screening; human.
XX Homo sapiens.
XX WO200009691-A2.
XX 24-FEB-2000.
XX 10-AUG-1999; 99WO-US018250.
XX PF 10-AUG-1998; 98US-0095982P.
XX PR 10-AUG-1998; 98US-0095982P.
XX (UROC-) UROGENESYS INC.
XX (AFAR/) AFAR D E.
XX (HUBE/) HUBERT R S.
XX (LEON/) LEONG K.
XX (RAIT/) RAITANO A B.
XX (SAFR/) SAFRAN D C.
XX (JAKO/) JAKOBOWITS A.
XX Afa DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
XX Jakobovits A;
XX WPI: 2000-206006/18.
XX N-PSDB; AA293040.
XX New isolated BPC-1 polypeptides, useful for developing products for the
XX diagnosis, staging, prognosis and treatment of cancers, particularly
XX prostate or bladder cancer.
XX Claim 1; Fig 1; 79PP; English.
XX BPC-1 polypeptides and polynucleotides can be used for the detection of
XX BPC-1 polypeptides and polynucleotides in biological samples, this is
XX particularly useful for detecting cancers expressing BPC-1, e.g. prostate
XX cancer or bladder cancer. Antibodies directed against BPC-1 or antisense
XX polynucleotides can be used for treating such cancers. The BPC-1

polypeptides can also be used in vaccines for treating or inhibiting the development of a cancer expressing Bpc-1. The polypeptides and polynucleotides can also be used for detection, prognosis, drug screening and predicting susceptibility to developing cancer. The Bpc-1 polypeptide comprises a CUB domain which is expressed in prostate and bladder carcinoma cells and which shows sequence similarity with CUB domains from other known proteins. In normal human tissues Bpc-1 is only expressed in certain tissues of the brain, however, it is expressed at high levels in prostate cancer cells and bladder cancer cells.

Sequence 158 AA:

Query Match 100.0%; Score 862; DB 3; Length 158;
Best Local Similarity 100.0%; Pred. No. 1,8e-90;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHGRSVLHIVASLIILHLSGATKKGTEKQTTSETOKSVQCGTWTGAEGGIFTSNPYS 60
DB 1 MHGRSVLHIVASLIILHLSGATKKGTEKQTTSETOKSVQCGTWTGAEGGIFTSNPYS 60
QY 61 KYPDPRECIYIIIAAPROCIELEFYDEKYSIEPSWECKFPHIEVRDGFSPPIIGRCGQ 120
DB 61 KYPDPRECIYIIIAAPROCIELEFYDEKYSIEPSWECKFPHIEVRDGFSPPIIGRCGQ 120
QY 121 QNPPVTKSSGRFLMIKFPADGELSMGFSARVNFPTP 158
DB 121 QNPPVTKSSGRFLMIKFPADGELSMGFSARVNFPTP 158

RESULT 2

ABU69144
ID ABU69144 standard; protein; 526 AA.

AC ABU69144;
XX 02-JUN-2003 (first entry)
XX

DE Human NOXV polypeptide #19.

XX Human; NOXV; metabolic disorder; diabetes; infectious disease; obesity;
KW anorexia; cancer; cardiovascular disorder; asthma; neurogenesis;
KW neurodegenerative disorder; epilepsy; immune disorder; osteoarthritis;
KW haematopoietic disorder; inflammatory skin disorder; dyslipidemia;
KW haematopoiesis; wound healing; angiogenesis; bacterial infection;
KW viral infection; fungal infection; helminthic infection; atherosclerosis;
KW protozoal infection; hypertension.

OS Homo sapiens.

XX WO200290504-A2.

XX 14-NOV-2002.

PF 02-MAY-2002; 2002MO-US014342.

XX 03-MAY-2001; 2001US-0286395P.
PR 04-MAY-2001; 2001US-0288980P.
PR 07-MAY-2001; 2001US-0289087P.
PR 14-MAY-2001; 2001US-0290753P.
PR 15-MAY-2001; 2001US-0291189P.
PR 16-MAY-2001; 2001US-0291243P.
PR 18-MAY-2001; 2001US-0292001P.
PR 21-MAY-2001; 2001US-0292374P.
PR 22-MAY-2001; 2001US-0292587P.
PR 23-MAY-2001; 2001US-0293107P.
PR 29-MAY-2001; 2001US-0294110P.
PR 30-MAY-2001; 2001US-0294434P.
PR 31-MAY-2001; 2001US-0294827P.
PR 18-JUN-2001; 2001US-0288988P.
PR 31-JUL-2001; 2001US-0308901P.
PR 17-AUG-2001; 2001US-0313388P.
PR 21-AUG-2001; 2001US-0313851P.
PR 21-AUG-2001; 2001US-0313937P.

PR 17-SEP-2001; 2001US-0322701P.
PR 17-SEP-2001; 2001US-0322802P.
PR 25-SEP-2001; 2001US-0324757P.
PR 27-SEP-2001; 2001US-0325314P.
PR 27-SEP-2001; 2001US-0325682P.
PR 21-NOV-2001; 2001US-0332129P.
PR 03-DEC-2001; 2001US-0336882P.
PR 14-DEC-2001; 2001US-0340305P.
PR 01-MAY-2002; 2002US-00138588.

XX (CURA-) CUREGEN CORP.

XX Alsobrook JP, Anderson DM, Boldog FL, Burgess CE, Casman SJ,
PI Chapoval A, Edinger S, Gerlach V, Gorman L, Gunther E, Guo X,
PI Kehoda R, Lepley DM, Li L, Liu X, Malayanhar UM, Miller CE,
PI Miller I, Padigaru M, Patnuranjan M, Pena CE, Rieger DK, Shenoy SG,
PI Shinkens RA, Spytek KA, Taupier RJ, Vernet CM, Voss EZ,
PI Zernusen BD;

XX MPI; 2003-103512/09.
DR N-PSDB; ACA10129.

XX New isolated NOXV polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOXV-associated disorders, e.g.
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.

PS Claim 2; Page 123; 340pp; English.

XX The invention relates to human NOXV polypeptides and the polynucleotides
XX encoding them. The polypeptides, polynucleotides and antibodies that bind
XX immunospecifically to the polypeptides are useful in the manufacture of a
XX medicament for treating a syndrome associated with a human disease,
XX preferably a NOXV-associated disorder. The sequences are useful for
XX treating, preventing or diagnosing diseases such as metabolic disorders,
XX diabetes, obesity, infectious diseases (viral, bacterial, fungal,
XX helminthic, and protozoal), anorexia, cancer, cardiovascular disorders
XX (e.g. hypertension, atherosclerosis), neurodegenerative disorders (e.g.
XX Alzheimer's disease, Parkinson's disease), epilepsy, immune disorders,
XX osteoarthritis, haematopoietic disorders, inflammatory skin disorders,
XX asthma and various dyslipidemias. The nucleic acids and polypeptides may
XX also be used as targets for the identification of small molecules that
XX modulate or inhibit e.g. neurogenesis, cell differentiation, cell
XX proliferation, haematopoiesis, wound healing and angiogenesis, and in the
XX generation of antibodies that bind immunospecifically to NOXV substances
XX for use in therapeutic or diagnostic methods. The nucleic acids are
XX further used as hybridisation probes, and in chromosome mapping, tissue
XX typing, preventive medicine and pharmacogenomics. Sequences ABU69126-
XX ABU69171 represent human NOXV polypeptides of the invention

XX Sequence 526 AA;

Query Match 98.7%; Score 851; DB 6; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.7e-88;
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHGRSVLHIVASLIILHLSGATKKGTEKQTTSETOKSVQCGTWTGAEGGIFTSNPYS 60

DB 1 MHGRSVLHIVASLIILHLSGATKKGTEKQTTSETOKSVQCGTWTGAEGGIFTSNPYS 60

QY 61 KYPDPRECIYIIIAAPROCIELEFYDEKYSIEPSWECKFPHIEVRDGFSPPIIGRCGQ 120

DB 61 KYPDPRECIYIIIAAPROCIELEFYDEKYSIEPSWECKFPHIEVRDGFSPPIIGRCGQ 120

QY 121 QNPPVTKSSGRFLMIKFPADGELSMGFSARVNFPTP 156

DB 121 QNPPVTKSSGRFLMIKFPADGELSMGFSARVNFPTP 156

RESULT 3

ABU79587
ID ABU79587 standard; protein; 533 AA.

XX

AC	AB079587;	
XX		
DT	21-OCT-2002 (first entry)	
XX		
DE	Human CUB domain-containing protein 39362.	
XX		
KW	CUB domain protein, 39362; human; antiarteriosclerotic; cardiant;	
KM	cerebroprotective; hypotensive; antirheumatic; antiarthritic;	
KW	antiallergic; neuroprotective; cytostatic; antithyroid; vasotrophic;	
KM	antialstematic; antidiabetic; nephrotropic; antiparkinsonian;	
KW	antiinflammatory; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Domain	1..344
FT		/note="predicted N-terminal extracellular domain"
FT	Peptide	1..23
FT		/label= Signal_peptide
FT	Modified-site	23..25
FT		/note="predicted protein kinase C phosphorylation site"
FT	Protein	24..533
FT		/label= Mature_protein
FT	Modified-site	24..27
FT		/note="predicted cAMP- and cGMP-dependent protein kinase phosphorylation site"
FT	Modified-site	27..29
FT		/note="predicted protein kinase C phosphorylation site"
FT	Modified-site	31..34
FT		/note="predicted casein kinase II phosphorylation site"
FT	Modified-site	35..37
FT		/note="predicted protein kinase C phosphorylation site"
FT	Domain	41..152
FT		/note="predicted CUB domain"
FT	Modified-site	50..55
FT		/note="predicted N-myristoylation site"
FT	Modified-site	129..131
FT		/note="predicted protein kinase C phosphorylation site"
FT	Modified-site	149..151
FT		/note="predicted protein kinase C phosphorylation site"
FT	Domain	172..284
FT		/note="predicted CUB domain"
FT	Modified-site	177..182
FT		/note="predicted N-myristoylation site"
FT	Modified-site	195..198
FT		/note="predicted casein kinase II phosphorylation site"
FT	Modified-site	241..244
FT		/note="predicted casein kinase II phosphorylation site"
FT	Modified-site	274..279
FT		/note="predicted N-myristoylation site"
FT	Modified-site	286..289
FT		/note="predicted casein kinase II phosphorylation site"
FT	Domain	290..328
FT		/note="predicted low density lipoprotein receptor class A domain"
FT	Modified-site	306..309
FT		/note="predicted N-glycosylation site"
FT	Modified-site	313..318
FT		/note="predicted N-myristoylation site"
FT	Modified-site	329..332
FT		/note="predicted cAMP- and cGMP-dependent protein kinase phosphorylation site"
FT	Modified-site	333..336
FT		/note="predicted casein kinase II phosphorylation site"
FT	Modified-site	340..343
FT		/note="predicted N-glycosylation site"
FT	Modified-site	341..351
FT		/note="predicted prokaryotic membrane lipoprotein lipid attachment site"
FT	Modified-site	343..348
FT		/note="predicted N-myristoylation site"
FT	Domain	345..363
FT		/note="predicted transmembrane domain"

FT	Domain	364. .533
FT	/note= "predicted C-terminal cytoplasmic domain"	
FT	Modified-site	377. .380
FT	/note= "predicted casein kinase II phosphorylation site"	
FT	Modified-site	397. .399
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	400. .405
FT	/note= "predicted N-myristoylation site"	
FT	Modified-site	421. .424
FT	/note= "predicted cAMP- and GMP-dependent protein kinase phosphorylation site"	
FT	Modified-site	424. .426
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	434. .439
FT	/note= "predicted N-myristoylation site"	
FT	Modified-site	439. .441
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	442. .447
FT	/note= "predicted N-myristoylation site"	
FT	Modified-site	446. .449
FT	/note= "predicted N-glycosylation site"	
FT	Modified-site	448. .451
FT	/note= "predicted casein kinase II phosphorylation site"	
FT	Modified-site	448. .450
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	481. .484
FT	/note= "predicted N-glycosylation site"	
FT	Modified-site	502. .504
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Modified-site	506. .509
FT	/note= "predicted casein kinase II phosphorylation site"	
FT	Modified-site	522. .525
FT	/note= "predicted casein kinase II phosphorylation site"	
FT	Modified-site	529. .532
FT	/note= "predicted N-glycosylation site"	
FT	Modified-site	530. .532
FT	/note= "predicted protein kinase C phosphorylation site"	
FT	Region	531. .533
PN	/note= "predicted microbody C-targeting signal"	
XX	WO200259275-A2.	
XX		
XX	01-AUG-2002.	
PD		
PF	08-JAN-2002; 2002WO-US000275.	
PR	08-JAN-2001; 2001US-0260286P.	
PA	(MILL-) MILLENNIUM PHARM INC.	
PI	Bandaru R;	
PJ		
XX	WPI, 2002-590825/63.	
DR	N-PSDB; ABRN84415.	
PT	New 39362 polypeptide and nucleic acid molecule, useful for detecting,	
PT	preventing or treating 39362-mediated or -related diseases, e.g.	
PT	atherosclerosis, cancer, and in screening assays, in predictive medicine	
PT	or pharmacogenetics.	
PS	Claim 5; Page 119; 131pp; English.	
XX		
CC	The present sequence is the protein sequence of 39362, a novel human CUB	
CC	domain-containing protein family member. The invention provides isolated	
CC	39362 proteins, fusion proteins, antigenic peptides and anti-39362	
CC	antibodies. It also provides isolated 39362 nucleic acid molecules,	
CC	antisense nucleic acid molecules, recombinant expression vectors, host	
CC	cells, and non-human transgenic animals, a method for identifying a	
CC	compound that binds 39362, a method for modulating 39362 polypeptide	
CC	activity, a method of inhibiting aberrant activity of a 39362-expressing	
CC	cell using a peptide, phosphopeptide, small organic molecule or molecule,	
CC	and a method of treating or preventing a disorder characterized by	
CC	aberrant activity of a 39362-expressing cell, especially a cardiovascular	

Claim 5; Page 119; 131pp; English.

The present sequence is the protein sequence of 39362, a novel human CUB domain-containing protein family member. The invention provides isolated 39362 proteins, fusion proteins, antigenic peptides and anti-39362 antibodies. It also provides isolated 39362 nucleic acid molecules, antisense nucleic acid molecules, recombinant expression vectors, host cells, and non-human transgenic animals, a method for identifying a compound that binds 39362, a method for modulating 39362 polypeptide activity, a method of inhibiting aberrant activity of a 39362-expressing cell using a peptide, phosphopeptide, small organic molecule or molecule, and a method of treating or preventing a disorder characterized by aberrant activity of a 39362-expressing cell, especially a cardiovascular

cell, by administering a compound that modulates the activity or expression of 39362 nucleic acid, 39362 proteins are useful in treating 39362-mediated or -related diseases, such as conditions involving cardiovascular disorders (e.g. atherosclerosis, myocardial infarction, aneurysm, stroke, hypertension), inflammatory disorders (e.g. rheumatoid arthritis, allergy, multiple sclerosis), and cellular proliferation or differentiation (e.g. cancer, Grave's disease, ischaemic disease). They may also be used in diagnosing, preventing or treating asthma, emphysema, chronic pulmonary disease, nephropathy, diabetes, hyperlipidemia, infections (e.g. bacterial, viral, parasitic), hepatic injury, Parkinson's disease, or Alzheimer's disease, anxiety or cardiovascular conditions associated with interventional procedures, such as restenosis following angioplasty, 39362 proteins, protein homologues, antibodies and nucleic acids are also useful as screening assays, predictive medicine (e.g. diagnostic assays, prognostic assays, forensic biology monitoring clinical trials, and pharmacogenetics), and methods of treatment (e.g. therapeutic and prophylactic).

Sequence 533 AA;

Query Match 98.7%; Score 851; DB 5; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.7e-88;
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MHGRSVLHIVASLIILHSGATKKTGKQTSETQKSVQCGTWTKHAGGIFTSPPNYS 60
1 MHGRSVLHIVASLIILHSGATKKTGKQTSETQKSVQCGTWTKHAGGIFTSPPNYS 60
QY 61 KYPDPRECTIIEAARQCIIELYPDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCQ 120
61 KYPDPRECTIIEAARQCIIELYPDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCQ 120
Db 121 QNPPIVTKSSGRPLWIKFADGELBSMGFSARYNFT 156
121 QNPPIVTKSSGRPLWIKFADGELBSMGFSARYNFT 156

RESULT 4

ADc42850
ID ADc42850 standard; protein; 525 AA.

ADc42850;
18-DEC-2003 (first entry)

REMAP protein #10.

Cytostatic; Antiarteriosclerotic; Anti-HIV; Antiinflammatory;
Anti-allergic; Antidiabetic; REMAP; pathogenesis.

Homo sapiens.

MO2003027228-A2.

03-APR-2003.

16-JUL-2002; 2002MO-US022833.

17-JUL-2001; 2001US-0306020P.
PR 27-JUL-2001; 2001US-0308179P.
PR 02-AUG-2001; 2001US-0309702P.
PR 10-AUG-2001; 2001US-0311476P.
PR 10-AUG-2001; 2001US-0311551P.
PR 10-AUG-2001; 2001US-0311718P.
PR 24-AUG-2001; 2001US-0314798P.
PR 31-AUG-2001; 2001US-0316639P.
PR 07-SEP-2001; 2001US-0317996P.

(INCYTE GENOMICS INC.

Lai PG, Honchel CD, Forsythe J, Walia NK, Tang TY, Borowsky M,
PI Barroso I, Yue H, Warren BA, Thangaveilu K, Gietzen KU, Asimzai Y,
Lee EA, Baughn MR, Gorvad AE, Dugan BM, Tran B, Li JX.

Richardson TW, Elliott VS, Zebajadian Y, Tran UK, Yao MG,
PI Peterson DP, Luo W, Lehr-Mason PM;
DR WPI; 2003-421156/39.
XX New human receptors and membrane-associated proteins (REMAP), useful for
PT diagnosing, treating or preventing disorders associated with aberrant
PT REMAP expression, e.g. cancer, AIDS, atherosclerosis, hypertension or
PT stroke.

Claim 1; SEQ ID NO 10; 115pp; English.

The present invention relates to an isolated polypeptide. The
XX polypeptides and polynucleotides are useful in diagnosing, treating and
XX preventing disorders associated with aberrant expression of REMAP, such
XX as cell proliferative, autoimmune/inflammatory, renal, neurological,
XX cardiovascular, metabolic, developmental, endocrine, muscle,
XX gastrointestinal, lipid metabolism or transport disorders, and viral
XX infections. These are also useful in assessing the effects of exogenous
XX compounds on the expression of nucleic acids and amino acid sequences of
XX REMAP, in facilitating drug discovery process, and in investigating the
XX pathogenesis of diseases or medical conditions. Expression and
XX purification were achieved using bacterial or virus-based expression
XX systems. The present sequence represents an REMAP protein of the
XX invention.

Sequence 525 AA;

Query Match 93.0%; Score 802; DB 7; Length 525;
Best Local Similarity 99.3%; Pred. No. 7e-83;
Matches 146; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 IVASLIILHSGATKKTGKQTSETQKSVQCGTWTKHAGGIFTSPPNYSKYPDPRECT 69
2 VVASLIILHSGATKKTGKQTSETQKSVQCGTWTKHAGGIFTSPPNYSKYPDPRECT 61
Db 70 YIIIEAARQCIIELYPDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCQ 129
62 YIIIEAARQCIIELYPDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCQ 121
QY 130 GRPLWIKFADGELBSMGFSARYNFT 156
122 GRPLWIKFADGELBSMGFSARYNFT 148

RESULT 5

ABBI0345
ID ABBI0345 standard; protein; 130 AA.

ABBI0345;

10-JAN-2002 (first entry)

Human cDNA SEQ ID NO: 653.

Human; gene therapy; neural disorder; immune system disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; cardiovascular disorder; renal disorder;
XX proliferative disorder; inflammation.

Homo sapiens.

MO200154474-A2.

02-AUG-2001.

17-JAN-2001; 2001MO-US001349.

31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.

CC DNAs. These can be used in the treatment of neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
CC is a protein of the invention
XX

SQ Sequence 130 AA;

Query Match 55.6%; Score 479; DB 4; Length 130;
Best Local Similarity 66.1%; Pred. No. 1.2e-46;
Matches 80; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 38 SVOCGTWTKHABGQIFTSNYPSPKYPDPDECIYIIIEAPROCIIELYDEKYSIEPSMECK 97
DB 9 ATQCGIWTWTSNGHFNASPNYDPSYPNKECIYIIIEAPRORIELTFDEHYIIEPSFECR 68
QY 98 FPHIEVRDGPFGFSPFIIGRCGQGNPVVTKSSGRFLMKIFPADGELSMGFSARXNFTPG 157
DB 69 FPHIEVRDGPFGFSPFIIDRYCGVKSPLLRSTGRFWMIKSSDELEGLGFRKXSFITG 128
QY 158 K 158
DB 129 K 129

RESULT 6
ABP66932
ID ABP66932 standard; protein; 130 AA.

XX ABP66932;
AC
XX
DT 09-DEC-2002 (first entry)
XX

DE Human polypeptide SEQ ID NO 653.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antihistaminatory; anti-HIV; antibacterial; antiviral;
KW antiparkinsonian; antischistosomal; antitubercular; cancer;
KW antihemetic; hepatoprotective; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; neurotropic; gene therapy; vaccine.

XX Homo sapiens.

XX US2002030672-A1.

XX 11-UTL-2002.

XX 17-JAN-2001; 2001US-00764853.

XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180688P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-UTL-2000; 2000US-0218390P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226688P.
PR 30-AUG-2000; 2000US-0228942P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.
PA (RUBEN/) RUBEN S M.
PA (BARA/) BARASH S C.

PI Rosen CA, Ruben SM, Barash SC;

DR MPI; 2002-681727/73.
DR N-PSDB; ABV83904.

XX Novel polypeptide useful for diagnosis, prognosis, prevention, and
PT treatment of immune, hyperproliferative, renal, respiratory,
PT cardiovascular, reproductive, endocrine, gastrointestinal and
PT neurological disorders.

XX Claim 11; SEQ ID NO 653; 369BP + Sequence Listing; English.

XX The invention relates to novel genes (ABV83682-ABV84101) and proteins
CC (ABP6710-ABP67129) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 130 AA;

Query Match 55.6%; Score 479; DB 5; Length 130;
Best Local Similarity 66.1%; Pred. No. 1.2e-46;
Matches 80; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 38 SVOCGTWTKHABGQIFTSNYPSPKYPDPDECIYIIIEAPROCIIELYDEKYSIEPSMECK 97
DB 9 ATQCGIWTWTSNGHFNASPNYDPSYPNKECIYIIIEAPRORIELTFDEHYIIEPSFECR 68
QY 98 FPHIEVRDGPFGFSPFIIGRCGQGNPVVTKSSGRFLMKIFPADGELSMGFSARXNFTPG 157

```
DB 69 FDHLEVRDGPFGFSLIRYGVKSPPLIRSTGRFMWIKFSSDELEGLGFPAKISFTPG 128
QY 158 K 158
Db 129 K 129

RESULT 7
AAU19919
ID AAU19919 standard; protein; 141 AA.
XX
AC AAU19919;
DT 04-DEC-2001 (first entry)
DE Novel human calcium-binding protein #28.
KW Human; calcium-binding protein; calcium flux; neurological disease;
KW immune dysfunction; digestive disorder; neoplastic disease;
KW blood disorder; infectious disease; gene therapy; immunosuppressive;
KW aneurysmal; cytostatic; vasotropic; antibacterial; nootropic;
KW virucide.
XX
OS Homo sapiens.
PN WO20015304-A2.
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001302.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190075P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0231415P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237940P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
```

PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-465568/50.
 N-PSDB; AAS31604.
 Isolated nucleic acid molecule encoding a calcium-binding protein is used in preventing, treating or ameliorating a medical condition.
 Claim 11: SEQ ID NO 116; 542bp; English.
 The present invention relates to the isolation of novel human calcium-binding proteins, and cDNA (AAS31577-AAS31654) and genomic sequences encoding for these proteins. The sequences of the invention are useful in the diagnosis, prevention and/or prognosis of diseases associated with aberrant calcium flux. Such disorders include neurological diseases (e.g. amyloidotic lateral sclerosis, ALS), immune dysfunction (e.g. severe combined immunodeficiency, SCID), digestive disorders (e.g. irritable bowel syndrome, IBS), neoplastic disease (e.g. cancer), blood disorders (e.g. haemophilia), and/or infectious disease (e.g. acquired immunodeficiency syndrome, AIDS). The novel calcium-binding proteins are also useful as screening tools to identify antagonists and/or agonists that may enhance or inhibit activities mediated by calcium-binding proteins. The polynucleotides of the invention are also useful in gene therapy. AAU19892-AAU19969 represent the novel human calcium-binding proteins. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 Sequence 141 AA;
 Query Match 55.6%; Score 479; DB 4; Length 141;
 Best Local Similarity 66.1%; Pred. No. 1,4e-46;
 Matches 80; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

AB011868
 ID AB011868 standard; peptide; 352 AA.
 AC AB011868;
 DT 11-JAN-2002 (first entry)
 DE Human secreted protein homologue, SEQ ID NO:2238.
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor; haematopoiesis regulation; tissue growth; immunomodulatory; activin; inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis; proliferation; metastasis; cancer; tumour; haematopoietic disorder; myeloid cell disorder; lymphoid cell disorder; asthma; arthritis; chronic inflammatory condition; proliferative retinopathy; atherosclerosis; coronary heart disease; arterial ischaemia; bone disorder; osteoporosis; vascular growth disorder; tissue regeneration; wound healing; infection; immune disorder; cell culture; drug screening; gene therapy; antiinflammatory; antiaustmatic; antiarthritic; haemostatic; antiarteriosclerotic; cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial; antifungal; vulnereary; anticulcer.
 OS Homo sapiens.
 PN WO200157188-A2.
 PD 09-AUG-2001.
 PF 05-FEB-2001; 2001WO-US003800.
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 (HYSR-) HYSRQ INC.
 Tang YT, Liu C, Drmanac RT;
 WPI; 2001-457740/49.
 N-PSDB; ABA09112.
 Human proteins and DNA encoding sequences useful for preventing, treating or ameliorating a medical condition in a mammalian subject e.g. arthritis and cancer.
 Claim 20; Page 268; 1963pp; English.
 Sequences AB019981-AB012330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides or antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound

XX Claim 8; SEQ ID NO 4135; 1380bp + Sequence Listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5' and 3' ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesized by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a polypeptide encoded by a full length
 CC human cDNA of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in CD-ROM
 CC format directly from EPO
 XX
 SQ Sequence 525 AA;
 Query Match 54.5%; Score 470; DB 4; Length 525;
 Best Local Similarity 54.9%; Pred. No. 8.8e-45;
 Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
 QY 6 SVLHI--VASLIILHLSGATKKGTSEKQTSSEKQSVQCGTWKHAEGIFTSPPNPSKYP 63
 DB 8 SVLKVLITLVVVEGIAVAKXTODGONIGIKHIPATCGIWRTSNGGHFASPNYDPSYP 67
 QY 64 PPRECIYITEAPROCIETLYFDEKYSIESWECKFPHIEVRDGFSPFIIGFCGQGNP 123
 DB 68 PKKECIYITEAPROCIETLYFDEHYYIESFEKRFPHLEVRDGFSPFIIDRYCGVSKP 127
 QY 124 PVKSGRFLMTKFPADGELESGFSARVNFEP 156
 DB 128 PLIRSTGRMWMIKFSDELEIGLGFRAKISFTIP 160
 RESULT 11
 AAU12228
 ID AAU12228 standard; protein; 525 AA.
 AC AAU12228;
 DT 24-OCT-2001 (first entry)
 XX Human PRO4401 polypeptide sequence.
 DE Human PRO4401 polypeptide sequence.
 XX Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;
 KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;
 KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;
 KW A-peptide; factor VIIA; gene therapy.
 OS Homo sapiens.
 PN WO20014046-A2.
 PD 07-JUN-2001.
 XX 01-DEC-2000; 2000WO-US032678.
 XX 01-DEC-1999; 99WO-US028301.
 XX 02-DEC-1999; 99WO-US028634.
 XX 02-DEC-1999; 99WO-US028651.
 XX 02-DEC-1999; 99WO-US028654.
 XX 02-DEC-1999; 99WO-US028655.
 XX 09-DEC-1999; 99US-0170282P.
 XX 16-DEC-1999; 99WO-US030095.
 XX 20-DEC-1999; 99WO-US030911.
 XX 20-DEC-1999; 99WO-US030999.
 XX 30-DEC-1999; 99WO-US031243.
 XX 30-DEC-1999; 99WO-US031274.
 XX 05-JAN-2000; 2000WO-US000219.
 XX 06-JAN-2000; 2000WO-US000277.
 XX 11-FEB-2000; 2000WO-US00376.
 XX 11-FEB-2000; 2000WO-US003565.

PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 03-MAR-2000; 2000US-0187202P.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 05-JUN-2000; 2000US-020832P.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 XX
 PA (GENENTECH INC.
 XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR N-PSDB; AAS21300.
 XX WPI: 2001-408281/43.
 DR Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing PRO
 PT polypeptides, and detect the presence of mammalian tumors e.g. lung,
 PT breast, prostate, cervical.
 XX
 PS Claim 12; Fig 114; 813bp; English.
 XX
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane PRO
 CC polypeptides. The PRO polypeptides are useful to detect other PRO
 CC polypeptides, to link bioactive molecules to cells expressing PRO
 CC polypeptides, to modulate biological activities of cells expressing PRO
 CC polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumors by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample. Some
 CC of the 275 sequences are also useful to stimulate the release of tumour
 CC necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or
 CC differentiation of chondrocytes, the proliferation or gene expression in
 CC pericyte cells, the release of proteoglycans from cartilage, the
 CC proliferation of inner ear utricular supporting cells or of T-
 CC lymphocytes, the release of a cytokine from peripheral blood monocytes
 CC (PBMCs), or the proliferation of endothelial cells. Some of the PRO
 CC polypeptides may modulate glucose or free fatty acid uptake by skeletal
 CC muscle cells or by adipocytes; or inhibit binding of A-peptide to factor
 CC VIIA. The PRO polypeptides can be used in assays to identify molecules
 CC involved in binding interactions. The polynucleotides encoding PRO
 CC transgenic or knock out animals and can be used in gene therapy
 CC
 XX
 SQ Sequence 525 AA;
 Query Match 54.5%; Score 470; DB 4; Length 525;
 Best Local Similarity 54.9%; Pred. No. 8.8e-45;
 Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
 QY 6 SVLHI--VASLIILHLSGATKKGTSEKQTSSEKQSVQCGTWKHAEGIFTSPPNPSKYP 63
 DB 8 SVLKVLITLVVVEGIAVAKXTODGONIGIKHIPATCGIWRTSNGGHFASPNYDPSYP 67
 QY 64 PPRECIYITEAPROCIETLYFDEKYSIESWECKFPHIEVRDGFSPFIIGFCGQGNP 123


```

Db      68 ENKECYIIIEAARQRIELTFDEHYIIESFECRPHLEVRGPGFSLIDRYGVSP 127
OY      124 PVIKSGRFMTKFPADGLESMSGFARINFTP 156
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      128 PLIRSTGRFMWIKFSDELEIGLGRKXYSFTP 160

RESULT 12
AAB47296
ID      AAB47296 standard, protein, 525 AA.
XX
XX      AAB47296;
AC
XX
DT      22-AUG-2001 (first entry)
XX
DE      PRO4401 polypeptide.
XX
KW      PRO; PRO1081; PRO1274; PRO10272; proliferation; T-lymphocyte; PRO1199;
KW      PRO1556; PRO4401; PRO10268; inhibition; stimulation; infiltration;
KW      mononuclear cell; eosinophil; erythema multiforme;
KW      polymorphonuclear neutrophil; PMN; antibody; immune-related disorder;
KW      systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
KW      juvenile chronic arthritis; spondyloarthropathy; systemic sclerosis;
KW      idiopathic inflammatory myopathy; Sjogren's syndrome; skin disease;
KW      systemic vasculitis; sarcoidosis; autoimmune haemolytic anaemia; asthma;
KW      autoimmune thrombocytopenia; thyroiditis; diabetes mellitus; allergy;
KW      immune-mediated renal disease; demyelination; central nervous system;
KW      peripheral nervous system; idiopathic demyelinating polyneuropathy;
KW      Guillain-Barre syndrome; hepatobiliary disease; eosinophilic pneumonia;
KW      chronic active hepatitis; primary biliary cirrhosis; allergic rhinitis;
KW      granulomatous hepatitis; sclerosing cholangitis; food hypersensitivity;
KW      inflammatory bowel disease; gluten-sensitive enteropathy; urticaria;
KW      Whipple's disease; idiopathic pulmonary fibrosis; contact dermatitis;
KW      psoriasis; atopic dermatitis; hypersensitivity pneumonitis;
KW      graft rejection; graft-versus-host disease.
XX
XX      Homo sapiens.
XX
FH      Key
FH      Peptide
FT      1. .22
FT      /label= Signal peptide
FT      22. .28
FT      /label= N-myristoylation site
FT      23. .525
FT      /label= Mature PRO4401
FT      32. .38
FT      /label= N-myristoylation site
FT      54. .60
FT      /label= N-myristoylation site
FT      186. .192
FT      /label= N-myristoylation site
FT      279. .285
FT      /label= N-myristoylation site
FT      311. .315
FT      /label= N-glycosylation site
FT      318. .324
FT      /label= N-myristoylation site
FT      348. .369
FT      /label= Transmembrane domain
FT      348. .354
FT      /label= N-myristoylation site
FT      352. .358
FT      /label= N-myristoylation site
FT      385. .389
FT      /label= N-glycosylation site
FT      426. .430
FT      /label= CAMP + CGMP-dependant protein kinase
FT      /label= phosphorylation site
FT      441. .447
FT      /label= N-myristoylation site
FT      453. .457
FT      /label= N-glycosylation site
FT

```

```

FT      Modified-site 475. .479
FT      /label= N-glycosylation site
FT
FT      Modified-site 479. .483
FT      /label= CAMP + CGMP-dependant protein kinase
FT      phosphorylation site
XX
XX      MO200140465-A2.
XX
PD      07-JUN-2001.
XX
PF      10-NOV-2000; 2000MO-US030873.
XX
XX      30-NOV-1999; 99MO-US028313.
XX      09-DEC-1999; 99US-0170262P.
XX      23-DEC-1999; 99US-0172059P.
XX      11-JAN-2000; 2000US-0175481P.
XX      20-JAN-2000; 2000US-0177118P.
XX      18-FEB-2000; 2000MO-US004342.
XX      03-MAR-2000; 2000US-0187202P.
XX      30-MAY-2000; 2000MO-US014941.
XX      05-JUN-2000; 2000US-0209832P.
XX      24-AUG-2000; 2000MO-US023328.
XX
XX      (GENE ) GENENTECH INC.
XX
PI      Forq S, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL;
PI      Hillan KJ, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX      WPI; 2001-381384/40.
DR      N-PSDB; AAC85966.
XX
XX      Isolated PRO polypeptide useful for treat or diagnose an immune-related
XX      disorder e.g. arthritis, asthma, allergy, diabetes or psoriasis.
XX
XX      Claim 1; Fig 12; 124pp; English.
XX
CC      The sequences given in AAB47291-99 show PRO polypeptides. PRO1081,
CC      PRO1274 and PRO10272 stimulate the proliferation of T-lymphocytes and
CC      PRO1199, PRO1556, PRO4401 and PRO10268 inhibit the proliferation of T-
CC      lymphocytes. PRO1754 and PRO0912 act to enhance the infiltration of
CC      mononuclear cells, eosinophils or polymorphonuclear neutrophils (PMN)
CC      into the tissue of a mammal. The PRO cDNA's and antibodies which bind to
CC      them, are used to treat an immune-related disorder in a mammal. Such
CC      disorders include systemic lupus erythematosus, rheumatoid arthritis,
CC      osteoarthritis, juvenile chronic arthritis, a spondyloarthropathy,
CC      systemic sclerosis, an idiopathic inflammatory myopathy, Sjogren's
CC      syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
CC      anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,
CC      immune-mediated renal disease, a demyelinating disease of the central or
CC      peripheral nervous system, idiopathic demyelinating polyneuropathy,
CC      Guillain-Barre syndrome, a chronic inflammatory demyelinating
CC      polyneuropathy, a hepatobiliary disease, infectious or autoimmune chronic
CC      active hepatitis, primary biliary cirrhosis, granulomatous hepatitis,
CC      sclerosing cholangitis, inflammatory bowel disease, gluten-sensitive
CC      enteropathy, Whipple's disease, an autoimmune or immune-mediated skin
CC      disease, a bullous skin disease, erythema multiforme, contact dermatitis,
CC      psoriasis, an allergic disease, asthma, allergic rhinitis, atopic
CC      dermatitis, food hypersensitivity, urticaria, an immunologic disease of
CC      the lung, eosinophilic pneumonia, idiopathic pulmonary fibrosis,
CC      hypersensitivity pneumonitis, a transplant rejection associated disease, graft
CC      rejection or graft-versus-host disease
XX
XX      Sequence 525 AA:
SQ

```

Query Match 54.5%; Score 470; DB 4; Length 525;
Best Local Similarity 54.9%; Pred. No. 8.8e-45;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

```

OY      6 SVLHI--VASLIILHSGATKGTETKQTSSEKSVQCGTWTKHAEGLFTSPNPKXP 63
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      8 SVLKVLITLVAVVEGIAVAQCKTDGQNTIGIKHPIPAQCGGIWRTSNQGHASPNYPSP 67
        |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY      64 PDKECYIIIEAARQRIELTFDEHYIIESFECRPHLEVRGPGFSLIDRYGVSP 123

```

Db 68 PKKECTYLLEAPRRIETLFDENRYIEBSPFCDHLEVRDGPFGSLIRYGVKSP 127
Qy 124 PVKSGRFLWIKFADGELSMGFSARYNFTP 156
Db 128 PLIRSTGRFMWIKFSSDELEBGLGRAKXSF 160

RESULT 13
ABOI7672
ID ABOI7672 standard; protein; 525 AA.
XX ABOI7672;
AC 26-AUG-2003 (first entry)
XX
XX
DE Novel human secreted and transmembrane protein PRO4401.
XX
KM Human; secreted and transmembrane protein; PRO; anti-inflammatory;
KM antiarteriosclerotic; cardiant; anti-fertility; anti-HIV; cytostatic;
KM antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;
KM TNF-alpha release; cell proliferation; cell differentiation;
KM gene expression modulator; proteoglycan release; cytokine release;
KM tumour; inflammatory disease; organ failure; atherosclerosis;
KM cardiac injury; infertility; birth defect; premature aging; AIDS;
KM acquired immunodeficiency syndrome; cancer; diabetic complication;
KM chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;
KM bioreactor; tissue typing.
OS Homo sapiens.
XX
XX US2003032156-A1.
XX
XX 13-FEB-2003.
XX
XX 06-MAY-2002; 2002US-00140474.
XX
XX 31-MAR-1997; 97WO-US005230.
XX 12-JUN-1998; 98WO-US012456.
XX 14-JUL-1998; 98WO-US014552.
XX 28-AUG-1998; 98WO-US017888.
XX 10-SEP-1998; 98WO-US018824.
XX 14-SEP-1998; 98WO-US019093.
XX 14-SEP-1998; 98WO-US019094.
XX 14-SEP-1998; 98WO-US019177.
XX 16-SEP-1998; 98WO-US019330.
XX 17-SEP-1998; 98WO-US019437.
XX 07-OCT-1998; 98WO-US021141.
XX 29-OCT-1998; 98WO-US022991.
XX 29-OCT-1998; 98WO-US022992.
XX 20-NOV-1998; 98WO-US024855.
XX 01-DEC-1998; 98WO-US025108.
XX 05-JAN-1999; 99WO-US000106.
XX 08-MAR-1999; 99WO-US005028.
XX 10-MAR-1999; 99WO-US005190.
XX 20-APR-1999; 99WO-US006615.
XX 14-MAY-1999; 99WO-US010733.
XX 02-JUN-1999; 99WO-US012252.
XX 01-SEP-1999; 99WO-US020111.
XX 08-SEP-1999; 99WO-US020594.
XX 13-SEP-1999; 99WO-US020944.
XX 15-SEP-1999; 99WO-US021090.
XX 15-SEP-1999; 99WO-US021547.
XX 15-SEP-1999; 99WO-US023089.
XX 05-OCT-1999; 99WO-US023089.
XX 29-NOV-1999; 99WO-US028214.
XX 30-NOV-1999; 99WO-US028313.
XX 01-DEC-1999; 99WO-US028409.
XX 01-DEC-1999; 99WO-US028301.
XX 01-DEC-1999; 99WO-US028634.
XX 02-DEC-1999; 99WO-US028551.
XX 02-DEC-1999; 99WO-US028564.
XX 02-DEC-1999; 99WO-US028565.
XX 16-DEC-1999; 99WO-US030095.

PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US020201.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00815744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 18-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00886342.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00887879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001US-00900827.
PR 18-JUL-2001; 2001US-00924419.
PR 06-AUG-2001; 2001US-00927796.
PR 09-AUG-2001; 2001US-00931836.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX
XX (GENTH) GENENTECH INC.
XX
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
XX Gierltsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,
XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI, 2003-341980/32.
XX N-PSDB; ACD23909.

PR 08-SEP-1999; 99WO-US020594.
 PR 13-SEP-1999; 99WO-US020944.
 PR 15-SEP-1999; 99WO-US021090.
 PR 15-SEP-1999; 99WO-US021547.
 PR 05-OCT-1999; 99WO-US023089.
 PR 29-NOV-1999; 99WO-US028214.
 PR 30-NOV-1999; 99WO-US028313.
 PR 30-NOV-1999; 99WO-US028409.
 PR 01-DEC-1999; 99WO-US028301.
 PR 01-DEC-1999; 99WO-US028634.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028564.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 20-DEC-1999; 99WO-US030911.
 PR 20-DEC-1999; 99WO-US030999.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004914.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005746.

(GETH) GENENTECH INC.

Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W, Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S, Smith V, Stewart TA, Tamas D, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-352836/33.
 N-PSDB; AC67050.

New isolated PRO polypeptide useful for treating diabetes, rheumatoid arthritis, sports injuries, obesity, hearing loss in mammals, stroke, or heart attack.

Claim 12; Fig 114; 643pp; English.

The present invention relates to the isolation of novel human PRO polypeptides, and the polynucleotide sequences encoding them. The PRO polypeptides and polynucleotides are useful for preparing a medicament useful in the treatment of diabetes, bone and/or cartilage disorders (e.g. rheumatoid arthritis, sports injuries, osteoarthritis), obesity, hyper- or hypo-insulinaemia, hearing loss, and coagulation disorders (e.g. stroke, heart attack). Anti-PRO antibodies are useful in diagnostic assays for PRO, by detecting its expression in specific cells, tissues or serum, and for affinity purification of PRO from recombinant cell culture or natural sources. AB08070-AB08114 represent the human PRO polypeptides of the invention. Note: The sequence data for this patent was obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov/psipdsidentry.html

Sequence 525 AA;

Query Match 54.5%; Score 470; DR 6; Length 525;
 Best Local Similarity 54.9%; Pred. No. 8.8e-45;
 Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

QY 6 SVLHI--VASLIIHLGATKKGKREKQTSVCCGWTAKAGGIFTSPPYKYP 63
 DB 8 SVLVLLITVVGGINAQAQTGQNGIKHIAATCGIIVKRSNGHFRASPPYPSYP 67
 QY 64 PDRECIYIEAAPROCIIEYDEKYSIEPSWECKFDHIEVADGPGFSPPIGRPCQOMP 123
 DB 68 PNRKCIYIEAAPROCIIEYDEKYSIEPSWECKFDHIEVADGPGFSPPIGRPCQOMP 127

QY 124 PVTKSGRFLMKPEADGELESMGFSARYNFTP 156
 DB 128 PLIRSTGRFMVTKSSDELEGLGRAXKSPFP 160

RESULT 15
 AB066626
 ID AB066626 standard; protein; 525 AA.

XX AB066626;

AC 23-MAY-2003 (first entry)

DT Human PRO polypeptide #57.

XX Human; PRO polypeptide; secreted and transmembrane protein;

KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;

KW differentiation; chondrocyte; tumour; genetic disorder; cytostatic.

XX Homo sapiens.

OS US2003036180-A1.

PN 20-FEB-2003.

PD 09-MAY-2002; 2002US-00143114.

XX 31-MAR-1997; 97WO-US005230.

PR 12-JUN-1998; 98WO-US012456.

PR 14-JUL-1998; 98WO-US014552.

PR 28-AUG-1998; 98WO-US017888.

PR 10-SEP-1998; 98WO-US018824.

PR 14-SEP-1998; 98WO-US019093.

PR 14-SEP-1998; 98WO-US019094.

PR 16-SEP-1998; 98WO-US019177.

PR 17-SEP-1998; 98WO-US019437.

PR 07-OCT-1998; 98WO-US021141.

PR 29-OCT-1998; 98WO-US022991.

PR 29-OCT-1998; 98WO-US022992.

PR 20-NOV-1998; 98WO-US024855.

PR 01-DEC-1998; 98WO-US025108.

PR 05-JAN-1999; 99WO-US000106.

PR 08-MAR-1999; 99WO-US005028.

PR 10-MAR-1999; 99WO-US005190.

PR 20-APR-1999; 99WO-US008615.

PR 14-MAY-1999; 99WO-US010733.

PR 02-JUN-1999; 99WO-US012252.

PR 01-SEP-1999; 99WO-US020111.

PR 08-SEP-1999; 99WO-US020594.

PR 13-SEP-1999; 99WO-US020944.

PR 15-SEP-1999; 99WO-US021090.

PR 15-SEP-1999; 99WO-US021547.

PR 05-OCT-1999; 99WO-US023089.

PR 29-NOV-1999; 99WO-US028214.

PR 30-NOV-1999; 99WO-US028313.

PR 30-NOV-1999; 99WO-US028409.

PR 01-DEC-1999; 99WO-US028301.

PR 01-DEC-1999; 99WO-US028634.

PR 02-DEC-1999; 99WO-US028551.

PR 02-DEC-1999; 99WO-US028564.

PR 02-DEC-1999; 99WO-US028565.

PR 16-DEC-1999; 99WO-US030095.

PR 20-DEC-1999; 99WO-US030911.

PR 20-DEC-1999; 99WO-US030999.

PR 22-DEC-1999; 99WO-US030720.

PR 30-DEC-1999; 99WO-US031243.

PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000277.

PR 06-JAN-2000; 2000WO-US000376.

PR 11-FEB-2000; 2000WO-US003565.

PR 18-FEB-2000; 2000WO-US004341.
 PR 18-FEB-2000; 2000WO-US004342.
 PR 22-FEB-2000; 2000WO-US004414.
 PR 24-FEB-2000; 2000WO-US004911.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005746.
 PR 10-MAR-2000; 2000WO-US005841.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 20-MAR-2000; 2000WO-US007377.
 PR 21-MAR-2000; 2000WO-US007532.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 11-AUG-2000; 2000WO-US022031.
 PR 23-AUG-2000; 2000WO-US023522.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 10-NOV-2000; 2000WO-US030873.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000WO-US034259.
 PR 20-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001US-0078498.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 01-MAR-2001; 2001WO-US006666.
 PR 09-MAR-2001; 2001US-00802706.
 PR 14-MAR-2001; 2001US-00808689.
 PR 22-MAR-2001; 2001US-00816744.
 PR 05-APR-2001; 2001US-00828366.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854280.
 PR 18-MAY-2001; 2001US-00860216.
 PR 25-MAY-2001; 2001US-00866028.
 PR 25-MAY-2001; 2001US-00866034.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001WO-US019682.
 PR 21-JUN-2001; 2001US-00887879.
 PR 22-JUN-2001; 2001WO-US020116.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 18-JUL-2001; 2001US-00908827.
 PR 06-AUG-2001; 2001US-00924419.
 PR 09-AUG-2001; 2001US-00927796.
 PR 16-AUG-2001; 2001US-00931836.
 PR 19-DEC-2001; 2001US-00028072.
 XX
 PA (GENENTECH INC.
 XX
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI; 2003-332040/31.
 DR N-PSDB; ACN03659.
 PT
 PT New secreted and transmembrane PRO nucleic acids, useful for gene
 PT therapy, in chromosome and gene mapping, as chromosome markers, in tissue
 PT typing, and in chromosome identification.
 XX
 PS Claim 12; Fig 114; 66opp; English.
 CC
 CC The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The PRO
 CC polypeptides are secreted and transmembrane proteins. The PRO

CC polypeptides are useful for detecting other PRO polypeptides, for linking
 CC bioactive molecules to cells expressing PRO polypeptides, for modulating
 CC biological activities of cells expressing PRO polypeptides, and for
 CC identifying agonists or antagonists. The PRO polypeptides are useful for
 CC for stimulating the release of tumour necrosis factor (TNF)-alpha from
 CC human blood, for stimulating the proliferation or differentiation of
 CC chondrocytes, and detecting the presence of tumours. The polynucleotide
 CC sequences encoding PRO polypeptides are useful as hybridisation probes,
 CC in chromosome and gene mapping, in the generation of antisense RNA and
 CC DNA, in the preparation of PRO polypeptides, for generating transgenic
 CC animals or knockout animals, for the genetic analysis of individuals with
 CC genetic disorders, and in gene therapy. AB06570-AB066844 represent the
 CC human PRO polypeptides of the invention. Note: The sequence data for this
 CC patent was obtained in electronic format directly from the USPTO web site
 CC at segdata.uspto.gov/pspsidentry.html
 CC
 XX

Sequence 525 AA;

Query Match 54.5%; Score 470; DB 6; Length 525;
 Best Local Similarity 54.9%; Pred. No. 8, 8e-45;
 Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

Qy 6 SVLHI--VASLIILHSGATKKGTEKQTTSERQKVQCGTWKHAEGIFTSNPNSKYP 63
 Db 8 SVLKVLITVLVVEGIAVAQKTDQNGINIKHIIPTQGIWRTSNGHFASPNPNDSPY 67
 Qy 64 PDRECIYIIIEAARQCIILYFDEKYSIEPSMECKPDHIEVRDGPFGSPFIIGRFGGQONP 123
 Db 68 PNKECIYIIIEAARQRIILTFDEHYIIEPSFECRDHIEVRDGPFGSPFIIDRYGVKSP 127
 Qy 124 PVKSSGFPLWIKFPADGELESMGFSARYNFTP 156
 Db 128 PLIRSTGRFWIKFSDELEGLGPRAKYSFTP 160

Search completed: May 18, 2004, 18:22:24
 Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: May 18, 2004, 18:19:46 ; Search time 39 Seconds

(without alignments)
1278.253 Million cell updates/sec

Title: US-09-887-593a-2

Perfect score: 862
Sequence: 1 MHGRSVLHIVASLIILHLS.....ADGELHSMGFSARYNTPPK 158

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	851	98.7	533	4	Q8TDF5	Q8TDF5 homo sapien
2	843	97.8	533	4	Q86W85	Q86W85 homo sapien
3	818	94.9	533	11	Q8R417	Q8R417 mus musculu
4	814	94.4	533	11	Q80X39	Q80X39 mus musculu
5	814	94.4	533	11	Q8C4S3	Q8C4S3 mus musculu
6	802	93.0	156	4	Q8TDF4	Q8TDF4 homo sapien
7	690	80.0	500	4	Q8ND78	Q8ND78 homo sapien
8	494	57.3	466	11	Q8NCM2	Q8NCM2 mus musculu
9	470	54.5	525	4	Q8NC67	Q8NC67 homo sapien
10	468	54.3	491	4	Q7Z381	Q7Z381 homo sapien
11	465	53.9	444	11	Q8C4O8	Q8C4O8 mus musculu
12	465	53.9	525	11	Q8BNC6	Q8BNC6 mus musculu
13	341	39.6	451	4	Q8ND51	Q8ND51 homo sapien
14	260	30.2	321	5	Q61849	Q61849 caenorhabdi
15	250	29.0	677	5	Q9VVC7	Q9VVC7 drosophila
16	238.5	27.7	1013	11	Q62381	Q62381 mus musculu

17	233.5	27.1	1008	13	Q9DER7	Q9DER7 gallus gall
18	233.5	27.1	1013	4	Q9NQ54	Q9NQ54 homo sapien
19	233.5	27.1	1013	4	Q43897	Q43897 homo sapien
20	229.5	26.6	574	5	Q86RL8	Q86RL8 illyanassa o
21	228.5	26.5	1012	11	Q9WVW6	Q9WVW6 mus musculu
22	227.5	26.4	1007	13	Q8J128	Q8J128 xenopus lae
23	227.5	26.4	3623	4	Q60494	Q60494 homo sapien
24	227.5	26.4	3623	11	Q70244	Q70244 rattus norv
25	227	26.3	926	4	Q9UQ00	Q9UQ00 homo sapien
26	227	26.3	977	13	Q91925	Q91925 xenopus lae
27	227	26.3	1015	4	Q9Y6L7	Q9Y6L7 homo sapien
28	226	26.2	735	13	Q57381	Q57381 xenopus lae
29	226	26.2	926	11	Q8QZT7	Q8QZT7 mus musculu
30	226	26.2	1019	13	Q57382	Q57382 xenopus lae
31	225	26.1	145	11	Q8BP20	Q8BP20 mus musculu
32	224.5	26.0	923	13	Q8OFX6	Q8OFX6 brachydanio
33	224.5	26.0	923	13	Q8AXP1	Q8AXP1 brachydanio
34	224	26.0	555	4	Q9H2E2	Q9H2E2 homo sapien
35	224	26.0	901	4	Q9H2E4	Q9H2E4 homo sapien
36	224	26.0	901	4	Q9H2D5	Q9H2D5 homo sapien
37	224	26.0	906	4	Q9H2D4	Q9H2D4 homo sapien
38	224	26.0	906	4	Q9H2E3	Q9H2E3 homo sapien
39	221	25.6	609	4	Q96190	Q96190 homo sapien
40	221	25.6	644	4	Q961H5	Q961H5 homo sapien
41	221	25.6	704	4	Q9H2E1	Q9H2E1 homo sapien
42	221	25.6	905	13	Q8U0L4	Q8U0L4 gallus gall
43	221	25.6	919	13	Q8UVR0	Q8UVR0 gallus gall
44	221	25.6	923	4	Q86T59	Q86T59 homo sapien
45	221	25.6	936	13	Q8UVQ9	Q8UVQ9 gallus gall

ALIGNMENTS

RESULT 1

Q8TDF5

AC Q8TDF5

DT 01-JUN-2002 (TREMURel. 21, Created)

DT 01-JUN-2002 (TREMURel. 21, Last sequence update)

DT 01-OCT-2003 (TREMURel. 25, Last annotation update)

OS Neopilin and tollid like-1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Retina;

RX MEDLINE=21940629; PubMed=11943477;

RA Stohr H., Berger C., Frohlich S., Weber B.H.;

RT "A novel gene encoding a putative transmembrane protein class A

RT extracellular CUB domains and a low-density lipoprotein class A

RT module: isolation of alternatively spliced isoforms in retina and

RT brain.";

RT Gene 286:223-231 (2002).

CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.

DR EMBL; AF448838; AAM18026.1; --

DR Genew; HGNC:13823; NEROL.

DR InterPro; IPR000859; CUB.

DR InterPro; IPR002172; LDL_receptor_A.

DR Pfam; PF00431; CUB; 2.

DR Pfam; PF00057; ldl_recept_a; 1.

DR SMART; SM00042; CUB; 2.

DR SMART; SM00192; LDLa; 1.

DR PROSITE; PS01180; CUB; 2.

DR PROSITE; PS50068; LDLRA_2; 1.

SQ SEQUENCE 533 AA; 60218 MW; 5B3BE7HBB8031B1 CRC64;

Query Match 98.7%; Score 851; DB 4; Length 533;

Best Local Similarity 100.0%; Pred. No. 1,le-85;

Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 M1HGRSVLHIVASLIILHLSGATKKGTEKQTTSSETOKSVQCGTWTAKHAGGIFTSNNYPS 60
DB 1 M1HGRSVLHIVASLIILHLSGATKKGTEKQTTSSETOKSVQCGTWTAKHAGGIFTSNNYPS 60
QY 61 KYPDPRECTIYIEAAROCIEIYFDEKYSIEPSWECKPFHIEVRDGPFGFSPIIGRFGCG 120
DB 61 KYPDPRECTIYIEAAROCIEIYFDEKYSIEPSWECKPFHIEVRDGPFGFSPIIGRFGCG 120
QY 121 QNPPVTKSGRFLMWIKFPADGELSMGFSARYNFTF 156
DB 121 QNPPVTKSGRFLMWIKFPADGELSMGFSARYNFTF 156

RESULT 2
ID Q86W85 PRELIMINARY; PRT; 533 AA.
AC Q86W85;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Neuropilin (NRP) and tollold (TLL)-like 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_Taxid=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC050329; AAH50329.1; -.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; ldl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
SQ SEQUENCE 533 AA; 60174 MW; DC1B7EBC2C3130E7 CRC64;

Query Match 97.8%; Score 843; DB 4; Length 533;
Best Local Similarity 99.4%; Pred. No. 8.3e-85;
Matches 155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 M1HGRSVLHIVASLIILHLSGATKKGTEKQTTSSETOKSVQCGTWTAKHAGGIFTSNNYPS 60
DB 1 M1HGRSVLHIVASLIILHLSGATKKGTEKQTTSSETOKSVQCGTWTAKHAGGIFTSNNYPS 60
QY 61 KYPDPRECTIYIEAAROCIEIYFDEKYSIEPSWECKPFHIEVRDGPFGFSPIIGRFGCG 120
DB 61 KYPDPRECTIYIEAAROCIEIYFDEKYSIEPSWECKPFHIEVRDGPFGFSPIIGRFGCG 120
QY 121 QNPPVTKSGRFLMWIKFPADGELSMGFSARYNFTF 156
DB 121 QNPPVTKSGRFLMWIKFPADGELSMGFSARYNFTF 156

RESULT 3
ID Q8R417 PRELIMINARY; PRT; 533 AA.
AC Q8R417;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Neuropilin and tollold like-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
RN (1)
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Retina;
RX MEDLINE=21940629; PubMed=11943477;
RA Stohr H., Berger C., Frohlich S., Weber B.H.;
RT "A novel gene encoding a putative transmembrane protein with two
RT extracellular CUB domains and a low-density lipoprotein class A
RT module; Isolation of alternatively spliced isoforms in retina and
RT brain.";
RL Gene 286:223-231(2002).
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AF448840; AAH18028.1; -.
DR MGI; MGI:2180216; Ncol.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; ldl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
SQ SEQUENCE 533 AA; 60214 MW; 6375B3972F8D558F CRC64;

Query Match 94.9%; Score 818; DB 11; Length 533;
Best Local Similarity 94.2%; Pred. No. 4.9e-82;
Matches 147; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 M1HGRSVLHIVASLIILHLSGATKKGTEKQTTSSETOKSVQCGTWTAKHAGGIFTSNNYPS 60
DB 1 M1HGRSVLHIVASLIILHLSGATKKGTEKQTTSSETOKSVQCGTWTAKHAGGIFTSNNYPS 60
QY 61 KYPDPRECTIYIEAAROCIEIYFDEKYSIEPSWECKPFHIEVRDGPFGFSPIIGRFGCG 120
DB 61 KYPDPRECTIYIEAAROCIEIYFDEKYSIEPSWECKPFHIEVRDGPFGFSPIIGRFGCG 120
QY 121 QNPPVTKSGRFLMWIKFPADGELSMGFSARYNFTF 156
DB 121 QNPPVTKSGRFLMWIKFPADGELSMGFSARYNFTF 156

RESULT 4
ID Q8OX39 PRELIMINARY; PRT; 533 AA.
AC Q8OX39;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Neuropilin-and tollold-like protein 1 (Brain-specific transmembrane
DE protein BTCL1).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toehlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Kodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smalhus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```


RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX PubMed:12810072;
 RA Michishita M., Ikeda T., Nakashiba T., Ogawa M., Tashiro K., Honjo T.,
 RA Doi K., Itohara S., Endo S.;
 RT "A novel gene, Brc1l, encoding CUB and LDL domains is expressed in
 RT restricted areas of mouse brain."
 RL Biochem. Biophys. Res. Commun. 306:680-686 (2003).
 DR EMBL: BC051145; AAH51145.1; -
 DR EMBL: AY138990; AAN38318.1; -
 DR InterPro: IPR000859; CUB;
 DR InterPro: IPR002172; LDL_receptor_A.
 DR Pfam: PF00431; CUB; 2.
 DR SMART: SM00057; ldl_recept_a; 1.
 DR SMART: SM00042; CUB; 2.
 DR SMART: SM00192; LDLa; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS50068; LDLRA_2; 1.
 DR Transmembrane.
 KM SEQUENCE 533 AA; 60242 MW; 6375B395421D558F CRC64;
 SQ
 Query Match 94.9%; Score 818; DB 11; Length 533;
 Best Local Similarity 94.2%; Pred. No. 4,9e-82;
 Matches 147; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MHHGSLVHIVASLIILHLSGATKKGTEKQTTSRTQKSVQCGTWTKAEGGIFTSPTNPS 60
 DB 1 MYGSLPHIILASLIILHSSGATKKGTGKQITPRTQKSVQCGTWTKAEGGVFTSPNPS 60
 QY 61 KYPDPRECIYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPFIIGRFGCG 120
 DB 61 KYPDPRECVYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPFIIGRFGCG 120
 QY 121 QNPPIKSSGRFLWIKFFADGELSMGFSARYNFTF 156
 DB 121 QNPPIKSSGRFLWIKFFADGELSMGFSARYNFTF 156
 RESULT 5
 Q8C4S3 PRELIMINARY; PRT; 533 AA.
 AC Q8C4S3;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical Microbodies C-terminal targeting signal/LDL-receptor
 DE class A.
 GN NETOL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCHI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL: AK081325; BAC38196.1; -
 DR MGD; MGI:2180216; Netol.
 DR InterPro: IPR000859; CUB.
 DR InterPro: IPR002172; LDL_receptor_A.

DR Pfam: PF00057; ldl_recept_a; 1.
 DR SMART: SM00042; CUB; 2.
 DR SMART: SM00192; LDLa; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS50068; LDLRA_2; 1.
 DR Hypothetical protein.
 KM SEQUENCE 533 AA; 60248 MW; 04319A6F9456CB3D CRC64;
 SQ
 Query Match 94.4%; Score 814; DB 11; Length 533;
 Best Local Similarity 93.6%; Pred. No. 1.4e-81;
 Matches 146; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MHHGSLVHIVASLIILHLSGATKKGTEKQTTSRTQKSVQCGTWTKAEGGIFTSPTNPS 60
 DB 1 MYGSLPHIILASLIILHSSGATKKGTGKQITPRTQKSVQCGTWTKAEGGVFTSPNPS 60
 QY 61 KYPDPRECIYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPFIIGRFGCG 120
 DB 61 KYPDPRECVYIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPFIIGRFGCG 120
 QY 121 QNPPIKSSGRFLWIKFFADGELSMGFSARYNFTF 156
 DB 121 QNPPIKSSGRFLWIKFFADGELSMGFSARYNFTF 156
 RESULT 6
 Q8TDF4 PRELIMINARY; PRT; 156 AA.
 AC Q8TDF4;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Neutrophilin and tollold like-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CX NCHI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RX MEDLINE=21940629; PubMed=11943477;
 RA Stroth H., Berger C., Frohlich S., Weber B.H.;
 RT extracellular CUB domains and a low-density lipoprotein class A
 RT module: isolation of alternatively spliced isoforms in retina and
 RT brain."
 RL Gene 286:223-231 (2002).
 CC -i- SIMILARITY: CONTAINS 1 CUB DOMAIN.
 DR EMBL: AF448839; AAM18027.1; -
 DR InterPro: IPR000859; CUB.
 DR Pfam: PF00431; CUB; 1.
 DR SMART: SM00042; CUB; 1.
 DR PROSITE: PS01180; CUB; 1.
 DR PROSITE: PS01180; CUB; 1.
 SQ SEQUENCE 156 AA; 17669 MW; 25F8B8F4553C2AB CRC64;
 Query Match 93.0%; Score 802; DB 4; Length 156;
 Best Local Similarity 99.3%; Pred. No. 7e-81;
 Matches 146; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 10 IVASLIILHLSGATKKGTEKQTTSRTQKSVQCGTWTKAEGGIFTSPTNPSKYPDPRECI 69
 DB 9 VASLIILHLSGATKKGTEKQTTSRTQKSVQCGTWTKAEGGIFTSPTNPSKYPDPRECI 68
 QY 70 YIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPFIIGRFGCGQNPPIKSS 129
 DB 69 YIIIEAPROCIIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPFIIGRFGCGQNPPIKSS 128
 QY 130 GRFLWIKFFADGELSMGFSARYNFTF 156
 DB 129 GRFLWIKFFADGELSMGFSARYNFTF 155

```

RESULT 7
Q8ND78      PRELIMINARY; PRT; 500 AA.
AC 08ND78;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP547B197.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Ansgorge W., Winkner U., Mewes H.W., Weil B., Wiemann S.;
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AL834354; CAD39019.1; -.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; Idl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
KM Hypothetical protein.
FT NON_TER
SQ SEQUENCE 500 AA; 56708 MW; 313628B650C49F08 CRC64;

Query Match
Best Local Similarity 80.0%; Score 690; DB 4; Length 500;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 ETOKSVQCGTWTXKHAAGGIFTSPNYSKYPPDRECIYIIIEAAROCIEIYFDEKYSIEPS 93
DB 1 ETOKSVQCGTWTXKHAAGGIFTSPNYSKYPPDRECIYIIIEAAROCIEIYFDEKYSIEPS 60

QY 94 WECKPDHIEVRDGPFGFSPITIGFCGQONPPVTKSSGRFLWIKFADGELISMGSFARVN 153
DB 61 WECKPDHIEVRDGPFGFSPITIGFCGQONPPVTKSSGRFLWIKFADGELISMGSFARVN 120

QY 154 FTP 156
DB 121 FTP 123

RESULT 8
Q8NCM2      PRELIMINARY; PRT; 466 AA.
AC 08NCM2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical Microbodies C-terminal targeting signal/LDL-receptor
DE class A (Fragment).
GN NETO1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RA MEDLINE=22354683; PubMed=1246681;
RA The FANTOM Consortium;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573 (2002).
DR EMBL; AK032510; BAC27902.1; -.

```

```

DR MGI:2180216; Netol.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; Idl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
KM Hypothetical protein.
FT NON_TER
SQ SEQUENCE 466 AA; 52840 MW; FE4AEA037EB344D9 CRC64;

Query Match
Best Local Similarity 57.3%; Score 494; DB 11; Length 466;
Matches 88; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 68 CYYIEAAROCIEIYFDEKYSIEPSWECKPDHIEVRDGPFGFSPITIGFCGQONPPVTK 127
DB 1 CYYIEAAROCIEIYFDEKYSIEPSWECKPDHIEVRDGPFGFSPITIGFCGQONPPVTK 60

QY 128 SSGRFLWIKFADGELISMGSFARVNFTP 156
DB 61 SSGRFLWIKFADGELISMGSFARVNFTP 89

RESULT 9
Q8NC67      PRELIMINARY; PRT; 525 AA.
AC 08NC67;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein FLJ90456.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuko Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Nimomiya K.;
RT "NEBO human cDNA sequencing project.";
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AK074937; BAC11303.1; -.
DR Genew; HGNC:14644; NETO2.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00057; Idl_recept_a; 1.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS50068; LDLRA_2; 1.
KM Hypothetical protein.
SQ SEQUENCE 525 AA; 59392 MW; EA6F98C3A88220EA CRC64;

Query Match
Best Local Similarity 54.5%; Score 470; DB 4; Length 525;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

QY 6 SVLHT--VASLIILHSGATKKGTSEKQTSFQKSVQCGTWTXKHAAGGIFTSPNYSKYTP 63
DB 8 SVLKVLLTLVWEGIAVAQKTDGQNIQIKHPAQCGIWTSTNGHFPASPNYDSDYP 67

QY 64 PDRECIYIIIEAAROCIEIYFDEKYSIEPSWECKPDHIEVRDGPFGFSPITIGFCGQONP 123
DB 68 PDRECIYIIIEAAROCIEIYFDEKYSIEPSWECKPDHIEVRDGPFGFSPITIGFCGQONP 127

```

QY 124 PVKSGRFLWIKFPADGELSMGFSARYNFTP 156
 DB 128 PLIRSTGRFMWIKFSSDBELGGLGFRAXKSFIP 160

RESULT 10

QY 10
 ID Q72381 PRELIMINARY; PRT; 491 AA.
 AC Q72381;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DE Hypothetical protein DKFZp686N19198 (Fragment).
 GN DKFZp686N19198.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Human fetal brain;
 RA Blocher H., Boeher M., Mewes H.W., Weil B., Amid C., Oanger A.,
 RA Fodor G., Han M., Wiemann S.;
 RU Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.
 DR EMBL; BX538063; CAD97994.1; -.
 KW Hypothetical protein.
 FT NON TER

QY 38 SVCCGTAKAEGGIFSPNPKYPPDRECIYITAAAPROCIELYPDEKYSIEPSWECK 97
 DB 8 ATQCGIWRISNGGHFASPNYPDSYPPKCEIYITAAAPROCIELTPEDEHYIIPSEFCR 67

Query Match 54.3%; Score 468; DB 4; Length 491;
 Best Local Similarity 65.5%; Pred. No. 3, 1e-43;
 Matches 78; Conservative 21; Mismatches 20; Indels 0; Gaps 0;

QY 98 FDHLEVRDGPFGSPITIGRFGGQONPPYKSSGRFLWIKFPADGELSMGFSARYNFTP 156
 DB 68 FDHLEVRDGPFGSPITIGRFGGQONPPYKSSGRFLWIKFPADGELSMGFSARYNFTP 126

RESULT 11

QY 08C408 PRELIMINARY; PRT; 444 AA.
 AC 08C408;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Hypothetical LDL-receptor class A (Fragment).
 GN NETO2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RA MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RT Nature 420:563-573(2002).
 DR EMBL; AK081462; BAC38225.1; -.
 DR MGD; MGI:1921763; NetO2.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR Pfam; PF000431; CUB; 2.
 DR Pfam; PF00057; ldl_recept_a; 1.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00192; LDLA; 1.
 DR PROSITE; PS01180; CUB; 2.

QY 10 IVASLIILHSGATKGTETKOTTSVQCGTWTAKAGGIFSPNPKYPPDRECI 69
 DB 14 LITVLVVEGIAVAQKTQCGQNGICHIIPATQCGIWRISNGGHFASPNYPDSYPPKCEI 73

QY 70 YITAAAPROCIELYPDEKYSIEPSWECKFDHLEVRDGPFGSPITIGRFGGQONPPYKSS 129
 DB 74 YITAAAPROCIELYPDEKYSIEPSWECKFDHLEVRDGPFGSPITIGRFGGQONPPYKSS 133

QY 130 GRFLWIKFPADGELSMGFSARYNFTP 156
 DB 134 GRFLWIKFPADGELSMGFSARYNFTP 160

QY 130 GRFLWIKFPADGELSMGFSARYNFTP 156
 DB 134 GRFLWIKFPADGELSMGFSARYNFTP 160

DR PROSITE; PSS0068; LDLA_2; 1.
 KW Hypothetical protein.
 FT NON TER
 SQ SEQUENCE 444 AA; 49350 MW; 4CB0A5C93C52657E CRC64;

Query Match 53.9%; Score 465; DB 11; Length 444;
 Best Local Similarity 53.7%; Pred. No. 6e-43;
 Matches 79; Conservative 31; Mismatches 37; Indels 0; Gaps 0;

QY 10 IVASLIILHSGATKGTETKOTTSVQCGTWTAKAGGIFSPNPKYPPDRECI 69
 DB 14 LITVLVVEGIAVAQKTQCGQNGICHIIPATQCGIWRISNGGHFASPNYPDSYPPKCEI 73

QY 70 YITAAAPROCIELYPDEKYSIEPSWECKFDHLEVRDGPFGSPITIGRFGGQONPPYKSS 129
 DB 74 YITAAAPROCIELYPDEKYSIEPSWECKFDHLEVRDGPFGSPITIGRFGGQONPPYKSS 133

QY 130 GRFLWIKFPADGELSMGFSARYNFTP 156
 DB 207 GRFLWIKFPADGELSMGFSARYNFTP 233

RESULT 12

QY 08BN06 PRELIMINARY; PRT; 525 AA.
 AC 08BN06;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Hypothetical LDL-receptor class A.
 GN NETO2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Body;
 RA MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium;
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RT Nature 420:563-573(2002).
 DR EMBL; AK083512; BAC38938.1; -.
 DR MGD; MGI:1921763; NetO2.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR Pfam; PF000431; CUB; 2.
 DR Pfam; PF00057; ldl_recept_a; 1.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00192; LDLA; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PSS0068; LDLA_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 525 AA; 59367 MW; 3E02308F311EC5CE CRC64;

QY 10 IVASLIILHSGATKGTETKOTTSVQCGTWTAKAGGIFSPNPKYPPDRECI 69
 DB 14 LITVLVVEGIAVAQKTQCGQNGICHIIPATQCGIWRISNGGHFASPNYPDSYPPKCEI 73

QY 70 YITAAAPROCIELYPDEKYSIEPSWECKFDHLEVRDGPFGSPITIGRFGGQONPPYKSS 129
 DB 74 YITAAAPROCIELYPDEKYSIEPSWECKFDHLEVRDGPFGSPITIGRFGGQONPPYKSS 133

QY 130 GRFLWIKFPADGELSMGFSARYNFTP 156
 DB 134 GRFLWIKFPADGELSMGFSARYNFTP 160

QY 130 GRFLWIKFPADGELSMGFSARYNFTP 156
 DB 134 GRFLWIKFPADGELSMGFSARYNFTP 160

RESULT 13

Q8ND51 PRELIMINARY; PRT; 451 AA.
 AC Q8ND51;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN DKFZP547L1153.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RL Bloeker H., Boecher M., Brandt P., Mewes H.W., Weil B., Wiemann S.;
 Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 DR EMBL; AL834404; CAD39066.1; -.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00057; LDL_receptor_a; 1.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM01192; LDL; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS00068; LDLRA_2; 1.
 DR PROSITE; PS01180; CUB; 2.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 451 AA; 51342 MW; 7BAD69517903B863 CRC64;
 Query Match 39.6%; Score 341; DB 4; Length 451;
 Best Local Similarity 68.6%; Pred. No. 3.5e-29;
 Matches 59; Conservative 16; Mismatches 11; Indels 0; Gaps 0;

QY 71 IIEAAPROCIETPEDEKYSIEPSWECKEPIHIVRSGPSPFPIGFCGQNPVTKSSG 130
 ID 061849 PRELIMINARY; PRT; 321 AA.
 AC 061849;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE K03E5.1 protein.
 GN K03E5.1
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodirinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RL MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 Gardner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,
 Jones M., Kersey J., Kirtlen J., Laister N., Latreille P.,
 Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
 Smalton N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;

RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans."
 RL Nature 368:32-38(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Becker M., Graves T., Fronick B.;
 RT "The sequence of C. elegans cosmid K03E5."
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 DR EMBL; AF067619; AAC17565.1; -.
 DR PIR; T33161; T33161.
 DR WormRep; K03E5.1; CE18007.
 DR InterPro; IPR000859; CUB.
 DR Pfam; PF00431; CUB; 2.
 DR SMART; SM00042; CUB; 2.
 DR PROSITE; PS01180; CUB; 2.
 SQ SEQUENCE 321 AA; 36995 MW; 1E2BDDFA70C3E36E CRC64;

Query Match 30.2%; Score 260; DB 5; Length 321;
 Best Local Similarity 43.5%; Pred. No. 2.3e-20;
 Matches 50; Conservative 20; Mismatches 37; Indels 8; Gaps 2;

QY 52 IFTSNYSKYPEDDEKCIYIEAAPROCIETPEDEKYSIEPSM-----ECKPHIEVR 104
 DB 75 IFTSNPDPYRPNDICVAVHSHRPHDHYVVFHFVHFHSTYDKIDAGEECNPFIER 134
 QY 105 DGPFGSPPIGRFCGQNP-VIKSGREPLIKFPADGELSMGFSARYNFTPGK 158
 DB 135 DRYGFSPLIAFCDDRMKREIRAVSGFLWIRFSDSMLTEYQFSAEVAIVPSK 189

RESULT 15

Q9VYC7 PRELIMINARY; PRT; 677 AA.
 AC Q9VYC7; Q8S2J8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE CG32635 protein (GH11189p).
 GN CG32635 OR CG12727 OR CG15751 OR CG15752.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RL MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-U., Andrews-Plambeck C., Baldwin D.,
 Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Borzhan M.R., Bouck J., Brodeur P., Brothier P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferriar C., Ferriar S., Fleischmann W.,
 Rosler C., Gabrielian A.E., Garg N.S., Gehart W.M., Glasser K.,
 Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

This Page Blank (uspto)

PRINTS: PR00020; MAMDOMAIN.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00231; FAS8C; 2.
 DR SMART; SM00137; MAM; 1.
 DR PROSITE; PS01180; CUB; 1.
 DR PROSITE; PS01285; FAS8C; 1; 2.
 DR PROSITE; PS01286; FAS8C; 2; 2.
 DR PROSITE; PS00022; FAS8C; 3; 2.
 DR PROSITE; PS00740; MAM; 1; 1.
 DR PROSITE; PS00060; MAM; 2; 1.
 DR Angiogenesis; Transmembrane; Glycoprotein; Neuron; Signal; Repeat;
 Receptor; Cell adhesion.
 KW SIGNAL
 FT CHAIN 1 18
 FT DOMAIN 19 914
 FT TRANSMEM 20 847
 FT DOMAIN 848 870
 FT DOMAIN 871 914
 FT DOMAIN 25 139
 FT DOMAIN 145 263
 FT DOMAIN 273 422
 FT DOMAIN 429 581
 FT DOMAIN 636 801
 FT DISULFID 25 52
 FT DISULFID 80 102
 FT DISULFID 145 171
 FT DISULFID 204 226
 FT DISULFID 273 422
 FT DISULFID 429 581
 SQ SEQUENCE 914 AA; 102480 MW; DD2EED6F0CB68C CRC64;
 Query Match 27.0%; Score 232.5; DB 1; Length 914;
 Best Local Similarity 35.3%; Pred. No. 4e-16; Mismatches 55; Indels 19; Gaps 4;
 Matches 54; Conservative 25; Mismatches 55; Indels 19; Gaps 4;
 QY 15 IILHSGATKTKTEKOTSETOKSVOCGWTAKBAGIFLTPSPYPSKYPPDRRCIYIIA 74
 DB 5 LFLHCAALT-----FTLSRALMSDKCGPTIKILSPGYLTPSPGYPSHPSCKEMLIOA 58
 QY 75 -APROCIETLYFDEKYSIEPSWECKPFHIEVRDPPFGSPILIGFCQGNPPVYKSSGRFL 133
 DB 59 PEYQIRIMINFNHFNPLEDR-DCKYDVEYIDGDNAGRLMKRYCGKIAPPLVSSGPFYL 117
 QY 134 WIKFADGELESWGFARSY-----NFT 155
 DB 118 FLKFVSDYETHGAGFSIRYEVFRGPECSNFT 150
 RESULT 2
 BML_XENIA STANDARD; PRT; 707 AA.
 ID BML_XENIA STANDARD; PRT; 707 AA.
 AC P98070;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Bone morphogenetic protein 1 precursor (EC 3.4.24.-) (BMP-1).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo; PubMed=8262384;
 RX MEDLINE=94085787; PubMed=8262384;
 RA Maeno M., Xue Y., Wood T.I., Ong R.C., Kung H.F.;
 RT "Cloning and expression of cDNA encoding Xenopus laevis bone
 morphogenetic protein-1 during early embryonic development.";
 RL Gene 134:257-261(1993).
 CC -1- FUNCTION: Involved in pattern formation in gastrula and later
 differentiation of developing organs.
 CC -1- DEVELOPMENTAL STAGE: Blastula, early gastrula and hatched
 tadpoles; little or no expression in morula and late gastrula.
 CC -1- SIMILARITY: Belongs to peptidase family M12A.

CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 3 CUB domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; J12249; AAA16313.1;
 DR PIR; JC2218; JC2218.
 DR HSSP; P00736; IAPQ.
 DR MEROPS; M12.005; -.
 DR InterPro; IPR000152; Asx_hydroxyl_S.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR006026; Peptidase_M.
 DR InterPro; IPR001506; Peptidase_M12A.
 DR Pfam; PF01400; Asclacin; 1.
 DR Pfam; PF00431; CUB; 3.
 DR Pfam; PF00008; EGF; 1.
 DR PRINTS; PR00480; ASTACIN.
 DR SMART; SM00042; CUB; 3.
 DR SMART; SM00179; EGF_CA; 1.
 DR SMART; SM00235; ZMGC; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS01180; CUB; 3.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00026; EGF_3; 1.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR Growth factor; Cytokine; Repeat; Osteogenesis; Chondrogenesis;
 KW Hydrolyase; Protease; Metalloprotease; EGF-like domain; Zinc; Calcium;
 KM Signal.
 FT PROPEP 1 2
 FT CHAIN 84 707
 FT DOMAIN 84 284
 FT DOMAIN 285 397
 FT DOMAIN 398 509
 FT DOMAIN 510 551
 FT DOMAIN 554 666
 FT METAL 176 176
 FT ACT SITE 177 177
 FT METAL 180 180
 FT METAL 186 186
 FT DISULFID 146 149
 FT DISULFID 514 526
 FT DISULFID 522 535
 FT DISULFID 537 550
 FT CARBOHYD 62 62
 FT CARBOHYD 105 105
 FT CARBOHYD 295 295
 FT CARBOHYD 326 326
 FT CARBOHYD 562 562
 SQ SEQUENCE 707 AA; 80673 MW; 1B6980D716DC98BD CRC64;
 Query Match 26.3%; Score 227; DB 1; Length 707;
 Best Local Similarity 40.2%; Pred. No. 1.1e-15;
 Matches 45; Conservative 18; Mismatches 47; Indels 2; Gaps 2;
 QY 41 CGWTAKBAGIFLTPSPYPSKYPPDRRCIYIIAAPROCIETLYFDEKYSIEPSWECKPFH 100
 DB 398 CGGEVKKDSCHI-QSPVYPPDYRPNKACWKLVSVEGFHGISF-QSFEIERHDSCAVDY 455
 QY 101 IEVRDPPFGSPILIGFCQGNPPVYKSSGRFLMIKFFADGELESWGFARSY 152
 DB 456 IEIRDSSETPVGRFCGYDKDPDIKSSITNQIKFVSDGSIKAGPSIANY 507


```

RESULT 3
NRP2_MOUSE
ID NRP2_MOUSE STANDARD, PRT; 931 AA.
AC Q35375; Q35373; Q35374; Q35376; Q35377; Q35378;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neuropilin-2 precursor (Vascular endothelial cell growth factor 165
  receptor 2).
GN NRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORMS A0; A17; A22; A5; B0 AND B5).
RC STRAIN=BALB/c;
RA Medline=97470888; PubMed=9311348;
RA Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;
RT "Neuropilin-2, a novel member of the neuropilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III."
RL Neuron 19:547-559 (1997).
CC -1- FUNCTION: High affinity receptor for semaphorins 3C, 3F, VEGF-165
CC and VEGF-145 isoforms of VEGF, and the PLGF-2 isoform of PGF.
CC -1- SUBUNIT: Neuropilin-2 probably forms a heteromeric complex with
CC neuropilin-1 in order to be a functional semaphorin E receptor.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Name=A22;
CC IsoId=Q35375-1; Sequence=Displayed;
CC Name=A0;
CC IsoId=Q35375-2; Sequence=VSP_004344;
CC Name=A5;
CC IsoId=Q35375-3; Sequence=VSP_004345;
CC Name=A17;
CC IsoId=Q35375-4; Sequence=VSP_004343;
CC Name=B0;
CC IsoId=Q35375-5; Sequence=VSP_004346;
CC Name=B5;
CC IsoId=Q35375-6; Sequence=VSP_004347;
CC -1- TISSUE SPECIFICITY: Expressed in developing CNS, PNS and in some
CC nonneural tissues including limb buds, developing bones, muscles,
CC intestinal epithelium, kidney, lung and submandibular gland.
CC -1- DEVELOPMENTAL STAGE: The expression pattern is very dynamic and is
CC developmentally regulated.
CC -1- SIMILARITY: Belongs to the neuropilin family.
CC -1- SIMILARITY: Contains 2 CUB domains.
CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
CC -1- SIMILARITY: Contains 1 MAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022866; AAC53379.1; -
DR EMBL; AF022864; AAC53377.1; -
DR EMBL; AF022865; AAC53378.1; -
DR EMBL; AF022857; AAC53380.1; -
DR EMBL; AF022858; AAC53381.1; -
DR EMBL; AF022861; AAC53382.1; -
DR HSSP; P12359; ICZT.
DR MGD; MGI:1100492; NRP2.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000421; FAS5_C.
DR InterPro; IPR008979; Gal_Bind_Like.

```

```

DR InterPro; IPRO00998; MAM_domain.
DR Pfam; PF00431; CUB_2.
DR Pfam; PF00754; F5_F8_type_C_2.
DR Pfam; PF00629; MAM; 1.
DR PRINTS; PRO0020; MAMDOMAIN.
DR SMART; SMO0042; CUB; 2.
DR SMART; SMO0231; FA5Bc; 2.
DR SMART; SMO0137; MAM; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FA5Bc_1; 2.
DR PROSITE; PS01286; FA5Bc_2; 2.
DR PROSITE; PS50022; FA5Bc_3; 2.
DR PROSITE; PS50060; MAM_2; 1.
KW Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Alternative splicing.
FT SIGNAL      1       20
FT CHAIN       21     931
FT DOMAIN      21     864
FT TRASMEN     865    889
FT DOMAIN      890    931
FT DOMAIN      28     142
FT DOMAIN      149    267
FT DOMAIN      277    427
FT DOMAIN      434    592
FT DOMAIN      642    802
FT DOMAIN      838    845
FT DISULFID    28      55
FT DISULFID    83     105
FT DISULFID   149     175
FT DISULFID   208     230
FT DISULFID    277    427
FT DISULFID    434    592
FT CAROHYD     152    152
FT CAROHYD     157    157
FT CAROHYD     629    629
FT CAROHYD     839    839
FT VARSPLIC    809    813
FT             /FTid=VSP 004343.
FT             Missing (in isoform A17).
FT VARSPLIC    809    830
FT             /FTid=VSP 004343.
FT             Missing (in isoform A0).
FT VARSPLIC    814    830
FT             /FTid=VSP 004344.
FT             Missing (in isoform A5).
FT VARSPLIC    810    931
FT             /FTid=VSP 004345.
FT             Missing (in isoform B0).
FT VARSPLIC    814    931
FT             /FTid=VSP 004346.
FT             Missing (in isoform B1).
FT FT          26.2%; Score 226; DB 1; Length 931;
FT FT          Best Local Similarity 38.1%; Pred. No. 1.9e-15;
FT FT          Matches 51; Conservative 18; Mismatches 51; Indels 14; Gaps 3.
SQ SEQUENCE    931 AA; 786 MW; 104558 MM; 76FD4343A11DD2F63 CRC64;

Query Match      26.2%; Score 226; DB 1; Length 931;
Best Local Similarity 38.1%; Pred. No. 1.9e-15;
Matches 51; Conservative 18; Mismatches 51; Indels 14; Gaps 3.

35 TOKSVGGTGWKHAAGGIFTSPPNPKPRPRECIIYIEA-APROCIELYVDDEKSTEPS 93
Db SQQDPDPCGRPNSSKDAGITISPGFYODPSHONCEWIVYAEPPOKVLANPHEIIE-K 80
94 WECKFHILEVADGFSGSPIIGRCGOQNPFVIKSSGRFLWIKEFPADGEISMGSARY- 152
DB HDCKYDFLEIRDDGESADILGKKCNIAPIPTLISSGVLYIKFTSDVARGAAGSLAYE 140
153 -----NFT 155
|||
```

Db 141 IFKTGSBDCSKNFT 154

RESULT 4

ID	NRP2_RAT	STANDARD	PRT	925 AA
AC	035276			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Neuropilin-2 precursor (Vascular endothelial cell growth factor 165 receptor 2).			
GN	NRP2			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_Taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley;			
RX	MEDLINE=97433085; PubMed=9288754;			
RA	Kolodkin A.L., Levengood D.V., Rowe E.G., Tai Y.-T., Giger R.J.,			
RA	Ginty D.D.;			
RT	"Neuropilin is a semaphorin III receptor.";			
RU	Cell 90:753-762(1997).			
CC	-1- FUNCTION: High affinity receptor for semaphorins 3C, 3F, VEGF-165 and VEGF-145 isoforms of VEGF, and the PLGF-2 isoform of PGF.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- TISSUE SPECIFICITY: Found in certain neuronal populations of the CNS and in other nonneuronal tissues including mesenchymal tissue lining in the ribs.			
CC	-1- SIMILARITY: Belongs to the neuropilin family.			
CC	-1- SIMILARITY: Contains 2 CUB domains.			
CC	-1- SIMILARITY: Contains 2 F5/8 type C domains.			
CC	-1- SIMILARITY: Contains 1 MAM domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	-----			
CC	EMBL: AF016297; AAC5338.1; -.			
DR	HSSP: P12259; ICZT.			
DR	InterPro: IPR000859; CUB.			
DR	InterPro: IPR000421; FAS8.C.			
DR	InterPro: IPR008979; Gal_bind_like.			
DR	InterPro: IPR000998; MAM_domain.			
DR	Pfam: PF00431; CUB_2.			
DR	Pfam: PF00754; F5_F8_type_C_2.			
DR	Pfam: PF00629; MAM_1.			
DR	PRINTS: PR00020; MAMDOMAIN.			
DR	SMART: SM00042; CUB; 2.			
DR	SMART: SM00231; FAS8C_2.			
DR	SMART: SM00137; MAM_1.			
DR	PROSITE: PS01180; CUB; 2.			
DR	PROSITE: PS01285; FAS8C_1; 2.			
DR	PROSITE: PS01286; FAS8C_2; 2.			
DR	PROSITE: PS50022; FAS8C_3; 2.			
DR	PROSITE: PS50060; MAM_2; 1.			
KW	Transmembrane; Glycoprotein; Neuromy; Signal; Repeat; Receptor.			
FT	SIGNAL	1	22	POTENTIAL.
FT	CHAIN	23	925	NEUROFILIN-2.
FT	DOMAIN	23	858	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	859	883	POTENTIAL.
FT	DOMAIN	884	925	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	28	142	CUB 1.
FT	DOMAIN	149	267	CUB 2.
FT	DOMAIN	277	427	F5/8 TYPE C 1.
FT	DOMAIN	434	592	F5/8 TYPE C 2.
FT	DOMAIN	642	802	MAM.

FT DISULFID 28 55 BY SIMILARITY.

FT DISULFID 83 105 BY SIMILARITY.

FT DISULFID 149 175 BY SIMILARITY.

FT DISULFID 208 230 BY SIMILARITY.

FT DISULFID 277 427 BY SIMILARITY.

FT DISULFID 434 592 BY SIMILARITY.

FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 833 833 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 834 834 N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 925 AA; 103896 MW; 3BF62903F64851C CRC64;

Query Match 26.1%; Score 225; DB 1; Length 925;

Best Local Similarity 37.3%; Pred. No. 2.4e-15;

Matches 50; Conservative 19; Mismatches 51; Indels 14; Gaps 3;

Db 141 IFKTGSBDCSKNFT 154

QY 153 -----NFT 155

Db 81 HDCKYDFLRDGSBSADLGSKHCNTAPFTISSGSVLYIKFTSDYARQAGFSLRYE 140

QY 35 TQSVQCGTWTGKAGGIFTSPPNPSKYPPDRECIYITBA-APROCIETLPDEKSIERS 93

Db 22 SQQDPCCGRUNSKDAGYITSPGYDYPSSHONCEWVYAPENOKIVAFNPFHEIE-K 80

QY 94 WECKEDHIEVRDGPFGPFIISRGCGQNPVYIKSGRFLMKFPADGELSEMGFSARY- 152

Db 81 HDCKYDFLRDGSBSADLGSKHCNTAPFTISSGSVLYIKFTSDYARQAGFSLRYE 140

QY 153 -----NFT 155

Db 141 IFKTGSBDCSKNFT 154

RESULT 5

ID	NRP2_HUMAN	STANDARD	PRT	931 AA
AC	060462; O14820; O14821;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Neuropilin-2 precursor (Vascular endothelial cell growth factor 165 receptor 2).			
GN	NRP2 OR VEGF165R2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM A0 AND A17).			
RX	MEDLINE=97470886; PubMed=9331348;			
RA	Chen H., Chedotal A., He Z.-G., Goodman C.S., Tessier-Lavigne M.;			
RT	"Neuropilin-2, a novel member of the neuropilin family, is a high affinity receptor for the semaphorins Sema E and Sema IV but not Sema III."			
RT	Neuron 19:547-559(1997).			
RU	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM A22).			
RC	TISSUE=Breast;			
RX	MEDLINE=98188099; PubMed=9529250;			
RA	Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;			
RT	"Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor.";			
RT	Cell 92:735-745(1998).			
RN	[3]			
RP	CHARACTERIZATION.			
RX	MEDLINE=20309748; PubMed=10748121;			
RA	Guzman-Polcorak Z., Cohen T., Herzog Y., Neufeld G.;			
RT	"Neuropilin-2 and neuropilin-1 are receptors of the 155-amino acid form of vascular endothelial growth factor (VEGF) and of placenta growth factor-2, but only neuropilin-2 functions as a receptor for the 145-amino acid form of VEGF.";			
RT	J. Biol. Chem. 275:18040-18045(2000).			
CC	-1- FUNCTION: High affinity receptor for semaphorins 3C, 3F, VEGF-165 and VEGF-145 isoforms of VEGF, and the PLGF-2 isoform of PGF.			
CC	-1- SUBUNIT: Neuropilin-2 probably forms a heteromeric complex with neuropilin-1 in order to be a functional semaphorin 3C receptor.			

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=A2;
 CC IsoId=O60462-1; Sequence=Displayed;
 CC Name=A0;
 CC IsoId=O60462-2; Sequence=VSP_004342;
 CC Name=A17;
 CC IsoId=O60462-3; Sequence=VSP_004341;
 CC -1- SIMILARITY: Belongs to the neuropilin family.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -1- SIMILARITY: Contains 1 MAM domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF022859; AAC51788.1; -
 CC EMBL; AF022860; AAC51789.1; -
 CC EMBL; AF016098; AAC12922.1; -
 CC HSSP; P12259; 1CZT
 CC Genew; HGNC:8005; NRP2.
 CC MIM; 602070; -
 CC GO; GO:0005624; C:membrane fraction; TAS.
 CC GO; GO:0004872; F:receptor activity; TAS.
 CC GO; GO:0005021; F:vascular endothelial growth factor receptor. . .; TAS.
 CC GO; GO:0007411; P:axon guidance; TAS.
 CC InterPro; IPR000859; CUB.
 CC InterPro; IPR000421; FAS8_C.
 CC InterPro; IPR008979; Gal Bind like.
 CC InterPro; IPR000998; MAM_domain.
 CC Pfam; PF00431; CUB; 2.
 CC Pfam; PF00754; F5_F8_type_C; 2.
 CC Pfam; PF00629; MAM; 1.
 CC PRINTS; PR00020; MAMDOMAIN.
 CC SMART; SMO0042; CUB; 2.
 CC SMART; SMO0231; FAS8C; 2.
 CC SMART; SMO0137; MAM; 1.
 CC PROSITE; PS01180; CUB; 2.
 CC PROSITE; PS01285; FAS8C_1; 2.
 CC PROSITE; PS01286; FAS8C_2; 2.
 CC PROSITE; PS50022; FAS8C_3; 2.
 CC PROSITE; PS50060; MAM_2; 1.
 CC Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
 KW Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 931
 FT DOMAIN 21 864
 FT TRANSEM 865 889
 FT DOMAIN 890 931
 FT DOMAIN 28 142
 FT DOMAIN 149 267
 FT DOMAIN 277 427
 FT DOMAIN 434 592
 FT DOMAIN 642 802
 FT DOMAIN 674 674
 FT DISULFID 28 55
 FT DISULFID 83 105
 FT DISULFID 149 175
 FT DISULFID 208 230
 FT DISULFID 277 427
 FT DISULFID 434 592
 FT CARBOHYD 152 152
 FT CARBOHYD 157 157
 FT CARBOHYD 629 629
 FT CARBOHYD 839 839
 FT VARSPIC 809 813

FT VARSPIC 809 830 Missing (in isoform A0).
 FT FT /FTId=VSP_004342.
 FT CONFLICT 602 602 E -> K (IN REF. 1).
 SQ SEQUENCE 931 AA; 104830 MW; 270CBAB69A0A797C CRC64;
 Query Match 26.0%; Score 224; DB 1; Length 931;
 Best local similarity 38.3%; Fred. No. 3.1e-15;
 Matches 51; Conservative 16; Mismatches 52; Indels 14; Gaps 3;
 QY 36 QKSVCGTWTGHAEGIFTSPNYPKPYPPDEKCIYIIFA-APROCIETLPDEKYSIPPSW 94
 DB 23 QPDPCCGRSLKSKAGYITSPGYPODYSHQNCENTVYAPEDNQTIVNPNFETIE-KH 81
 QY 95 ECKEDHIEVRDPPFGFPDIIIRFCGQNPPIYKSSGRPLWIKFFADGELSMGFSARY-- 152
 DB 82 DCKDVFIRIDGDSHADLLKHGQCNINAPPTIISGSMWLYIKFTSDYARQGGFSLRYEI 141
 QY 153 -----NFT 155
 DB 142 FKTSQSDCSKNFT 154
 RESULT 6
 EMBL_MOUSE
 ID EMBL_MOUSE STANDARD; PRT; 991 AA.
 AC P98063;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Bone morphogenetic protein 1 precursor (BC 3.4.24.19) (BMP-1)
 DE (Procollagen C-proteinase) (PCP) (Mammalian tollid protein) (mtId).
 GN BMP1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Embryo;
 RX MEDLINE=94229342; PubMed=8174772;
 RA Fukagawa M., Noboru S., Hogan B.L.M., Jones C.M.;
 RT "Embryonic expression of mouse bone morphogenetic protein-1 (BMP-1),
 RT which is related to the Drosophila dorsventral gene tollid and
 RT encodes a putative astacin metalloendopeptidase.";
 RL Dev. Biol. 163:175-183(1994).
 CC -1- FUNCTION: Cleaves the C-terminal propeptides of procollagen I, II
 CC and III. Induces cartilage and bone formation.
 CC -1- CATALYTIC ACTIVITY: Cleavage of the C-terminal propeptide at
 CC Ala-I-Asp in type I and II procollagens and at Arg-I-Asp in type
 CC III.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -1- ENZYME REGULATION: Activity is increased by the procollagen C-
 CC endopeptidase enhancer protein.
 CC -1- TISSUE SPECIFICITY: At high levels in embryonic maternal deciduum
 CC and floor plate region of the neural tube. Less in developing
 CC membranous and endochondral bone, submucosa of intestine, dermis
 CC of skin and the mesenchyme of spleen and lung.
 CC -1- SIMILARITY: Belongs to peptidase family M12A.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 5 CUB domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L24755; AAA37306.1; -
 CC PTR; I49540; I49540.
 DR HSSP; P00736; IAPQ.
 DR MEROPS; M12.005; -.

CC IsoId=014786-2; Sequence=VSP_004339, VSP_004340;
 CC -1- TISSUE SPECIFICITY: The expression of isoforms 1 and 2 does not
 CC seem to overlap. Isoform 1 is expressed by the blood vessels of
 CC different tissues. In the developing embryo it is found
 CC predominantly in the nervous system. In adult tissues, it is
 CC highly expressed in heart and placenta; moderately in lung, liver,
 CC skeletal muscle, kidney and pancreas; and low in adult brain.
 CC Isoform 2 is found in liver hepatocytes, kidney distal and
 CC proximal tubules.
 CC -1- SIMILARITY: Belongs to the neuropilin family.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC -1- SIMILARITY: Contains 2 F5/F8 type C domains.
 CC -1- SIMILARITY: Contains 1 MAM domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF018956; AAC51759.1; -
 CC EMBL; AF016050; AAC12921.1; -
 CC EMBL; AF145712; AAF44344.1; -
 CC PDB; 1KEX; 28-JAN-03.
 CC Genew; HGNC:8004; NRPL.
 CC MIM; 602069; -
 CC DR GO; GO:0005021; F:vascular endothelial growth factor receptor. . . ; TAS.
 CC DR GO; GO:0007411; P:axon guidance; TAS.
 CC DR GO; GO:0007267; P:cell-cell signaling; TAS.
 CC DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 CC DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
 CC DR GO; GO:0007165; P:signal transduction; TAS.
 CC DR InterPro; IPR000859; CUB.
 CC DR InterPro; IPR000421; FAS8_C.
 CC DR InterPro; IPR008979; Gal bind like.
 CC DR InterPro; IPR000998; MAM_domain.
 CC DR Pfam; PF00431; CUB; 2.
 CC DR Pfam; PF00754; F5_F8_type_C; 2.
 CC DR Pfam; PF00629; MAM; 1.
 CC DR PRINTS; PR00020; MAMDOMAIN.
 CC DR SMART; SM00042; CUB; 2.
 CC DR SMART; SM00231; FAS8C; 2.
 CC DR SMART; SM00137; MAM; 1.
 CC DR PROSITE; PS01180; CUB; 2.
 CC DR PROSITE; PS01285; FAS8C_1; 2.
 CC DR PROSITE; PS01286; FAS8C_2; 2.
 CC DR PROSITE; PS00022; FAS8C_3; 2.
 CC DR PROSITE; PS00740; MAM_1; 1.
 CC DR PROSITE; PS00060; MAM_2; 1.
 CC DR Angiogenesis; Transmembrane; Glycoprotein; Neuron; Signal; Repeat;
 CC Receptor; Alternative splicing; 3D-structure.
 CC KM
 CC PT SIGNAL 1 21
 CC FT CHAIN 22 923 NEURAPILIN-1.
 CC FT DOMAIN 22 856 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 857 879 POTENTIAL.
 CC FT DOMAIN 880 923 CYTOPLASMIC (POTENTIAL).
 CC FT DOMAIN 27 141 CUB 1.
 CC FT DOMAIN 147 265 CUB 2.
 CC FT DOMAIN 275 424 F5/F8 TYPE C 1.
 CC FT DOMAIN 431 583 F5/F8 TYPE C 2.
 CC FT DOMAIN 645 811 MAM.
 CC FT DISULFID 27 54 PROBABLE.
 CC FT DISULFID 82 104 PROBABLE.
 CC FT DISULFID 147 173 PROBABLE.
 CC FT DISULFID 206 228 PROBABLE.
 CC FT DISULFID 275 424 BY SIMILARITY.
 CC FT DISULFID 431 583 BY SIMILARITY.
 CC FT CARBOHYD 150 583 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VASAPLIC 642 644 EEP -> GIK (in isoform 2).
 FT VASAPLIC 645 923 /FtId=VSP_004339.
 FT VASAPLIC 645 923 Missing (in isoform 2).
 FT CONFLICT 26 26 /FtId=VSP_004340.
 FT CONFLICT 749 749 K -> E (IN REF. 1).
 FT CONFLICT 855 855 D -> H (IN REF. 2).
 FT CONFLICT 855 855 E -> D (IN REF. 2).
 SQ SEQUENCE 923 AA; 103120 MW; ADEADCA484955D57 CRC64;
 Query Match 25.6%; Score 221; DB 1; Length 923;
 Best local similarity 41.2%; Pred. No. 6.3e-15;
 Matches 47; Conservative 18; Mismatches 47; Indels 2; Gaps 2;
 Oy 40 QCGTWKAEIGGITSPVPSKPPDRECIYIEA-PROCTIEYDEKSIPEWCKE 98
 Db 26 KCGDTIKESGYLSPGYHPHSYSEKCEWLQAPDPYQIMNFNHPLEDR-DCKY 84
 Oy 99 DHIYRQGFSPSPILGRPCGQNPVTKSSGRFLWIKFPADGEISMGFSARY 152
 Db 85 DYVEVFGENENGHFRGKFCGKIAPPVVS8GPELFIFKIVSDYETHGAGFSIRY 138
 RESULT 8
 BMD1 HUMAN STANDARD; PRT; 986 AA.
 ID BMD1 HUMAN
 AC P13497; Q13292; Q13872; Q14874; Q99421; Q99422; Q99423; Q9UL38;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Bone morphogenetic protein 1 precursor (EC 3.4.24.19) (BMP-1)
 DE (Procollagen C-proteinase) (PCP) (Mammalian tollold protein) (MtlD).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP TISSUE=FROM N.A. (ISOFORM BMD1-3).
 RC TISSUE=Skin;
 RX MEDLINE=96209868; PubMed=6643539;
 RA Li S.W., Sieron A.L., Ferrata A., Hojima Y., Arnold W.V.,
 RA Prockop D.J.;
 RT "The C-proteinase that processes procollagens to fibrillar collagens
 RT is identical to the protein previously identified as bone morphogenic
 RT protein-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:5127-5130(1996).
 RP SEQUENCE FROM N.A. (ISOFORM BMD1-1).
 RX MEDLINE=89072730; PubMed=3201241;
 RA Wozney J.M., Rosen V., Celeste A.J., Mittlek L.M., Whitters M.J.,
 RA Kriz R.W., Hewick R.M., Wang E.A.;
 RT "Novel regulators of bone formation: molecular clones and
 RT activities";
 RL Science 242:1528-1534(1988).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS BMD1-4; BMD1-5 AND BMD1-6).
 RC TISSUE=Placenta;
 RX MEDLINE=98160316; PubMed=9500680;
 RA Janitz M., Heiser V., Boettcher U., Landt O., Lauster R.;
 RT "Three alternatively spliced variants of the gene coding for the human
 RT bone morphogenetic protein-1.";
 RL J. Mol. Med. 76:141-146(1998).
 RN [4]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS BMD1-3 AND BMD1-7).
 RC TISSUE=Placenta;
 RX MEDLINE=95096114; PubMed=7796260;
 RA Takahara K., Lyons G.E., Greenspan D.S.;
 RT "Bone morphogenetic protein-1 and a mammalian tollold homologue (mtlD)
 RT are encoded by alternatively spliced transcripts which are
 RT differentially expressed in some tissues.";
 RL J. Biol. Chem. 269:32572-32578(1994).
 RN [5]

FT VARSPLIC 824 986 /FTID=VSP 005469.
FT Missing (in isoform BMP1-7).
FT CONFLICT 748 748 /FTID=VSP_005470.
FT CONFLICT 934 934 D -> N (IN REF. 4).
FT CONFLICT 934 934 R -> S (IN REF. 4).
SQ SEQUENCE 986 AA; 111248 MM; F89201913AC3CBEA CRC64;
Query Match 25.6%; Score 220.5; DB 1; Length 986;
Best Local Similarity 38.2%; Pred. No. 7,6e-15;
Matches 39; Conservative 24; Mismatches 38; Indels 1; Gaps 1;
OY 51 GIFTSPNYSKYPRECITYIIEAPROCIELYPDEKYSIPEWCKEDHIEVRDGPGRF 110
DB 756 GFTSPNPNPKKPKSKKCECTMAISSIFGRVKLTME-MDISQPCAYDHLVFDGRDAK 814
OY 111 SPITRFGQONNPVTKSGRFLWIKFPADGELMSGFSARY 152
DB 815 APVLGRFCGSKKPEVLATGSRMFLRFYSNDSNVCKGQASH 856
RESULT 9
TID BRARE STANDARD; PRT; 1022 AA.
AC 057460;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dorsal-ventral patterning tolloid protein precursor (EC 3.4.24.-)
DE (Mini fin protein).
GN TOLLID OR TLD OR MPN.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Gastrula.
RX MEDLINE=98057457; PubMed=9395394;
RA Blader P., Rastegar S., Fischer N., Straehle U.,
RT "Cleavage of the BMP-4 antagonist chordin by zebrafish Tolloid.";
RL Science 278:1937-1940(1997).
RN [2]
RP FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=99307076; PubMed=10375503;
RA Connors S.A., Trout J., Ekker M., Mullins M.C.;
RT "The role of tolloid/mini fin in dorsoventral pattern formation of the
RT zebrafish embryo.";
RL Development 126:3119-3130(1999).
CC -|- FUNCTION: Required for patterning ventral tissues of the tail. May
CC increase bone morphogenetic protein (BMP) activity at the end of
CC gastrulation by proteolytic cleavage of chordin and release of BMP
CC from inactive complexes.
CC -|- TISSUE SPECIFICITY: During gastrulation, accumulates around the
CC closing blastopore with greater expression ventrally. At the
CC animal pole, expressed in the ectoderm flanking the anterior
CC neural plate. At the 10-somite stage, expressed in the developing
CC tailbud and cranial neural crest. At the 20-somite stage, also
CC expressed in the hematopoietic system.
CC -|- SIMILARITY: Belongs to peptidase family M2A.
CC -|- SIMILARITY: Contains 2 EGF-like domains.
CC -|- SIMILARITY: Contains 5 CUB domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; AF027596; AAC60304.1; -;
DR HSSP; P35555; IEMN.

DR MEROPS; M12.016; -;
DR ZFIN; ZDB-GENE-990415-265; tolloid.
DR InterPro; IPR000152; Asx_hydroxyl_s.
DR InterPro; IPR000859; CUB-
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006025; Pept_M_zn_BS.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001506; Peptidase_M12A.
DR Pfam; PF01400; Asclacin; 1.
DR Pfam; PF00431; CUB; 5.
DR Pfam; PF00008; EGF; 2.
DR PRINTS; PRO0480; ASTACIN.
DR SMART; SM00042; CUB; 5.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00235; ZmC; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01180; CUB; 5.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Developmental protein; Hydrolase; Protease; Metalloprotease; Zinc;
KW Metal-binding; Calcium; EGF-like domain; Repeat; Signal; Glycoprotein;
KW Zymogen.
FT SIGNAL 1 32
FT PROPEP 33 156
FT CHAIN 157 1022
FT DOMAIN 157 357
FT DOMAIN 358 470
FT DOMAIN 471 583
FT DOMAIN 583 624
FT DOMAIN 627 739
FT DOMAIN 739 779
FT DOMAIN 783 895
FT METAL 896 1012
FT ACT_SITE 249 249
FT METAL 250 250
FT METAL 253 253
FT METAL 259 259
FT DISULFID 358 384
FT DISULFID 411 433
FT DISULFID 471 497
FT DISULFID 524 546
FT DISULFID 587 599
FT DISULFID 595 608
FT DISULFID 610 623
FT DISULFID 627 653
FT DISULFID 680 702
FT DISULFID 743 754
FT DISULFID 750 763
FT DISULFID 765 778
FT DISULFID 783 809
FT DISULFID 836 858
FT DISULFID 896 926
FT DISULFID 953 975
FT CARBOHYD 129 129
FT CARBOHYD 178 178
FT CARBOHYD 368 368
FT CARBOHYD 399 399
FT CARBOHYD 635 635
SQ SEQUENCE 1022 AA; 115536 MM; A68CAID0E411793F9 CRC64;
Query Match 25.6%; Score 220.5; DB 1; Length 1022;
Best Local Similarity 40.4%; Pred. No. 7,9e-15;
Matches 42; Conservative 19; Mismatches 42; Indels 1; Gaps 1;
OY 49 EGITFSPNYSKYPRECITYIIEAPROCIELYPDEKYSIPEWCKEDHIEVRDGPGRF 108

DB 478 DSGQIQSNYPDDYDRPSKCEWIRITVSGYVGLSF-QVFELIHRDSCAVDYLEVRDGLS 536
 QY 109 GSPPIRGFCGQONPPVITKSGRFLMTFRPADGLESGFGRY 152
 DB 537 ENSPIRGFCGQKEDIRSTSNLMKRVSDGTVMKAGFANF 580

RESULT 10
 TLD_DROME STANDARD; PRT; 1057 AA.

AC P25723; Q9VC46;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dorsal-ventral patterning tolloid protein precursor (BC 3.4.24.-).
 GN TLD OR CG6868.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBITaxID=7227;
 RN (1)
 RN SEQUENCE FROM N.A.
 RC STRAIN=Canlon-S;
 RX MEDLINE=92034970; PubMed=1840509;
 RA Shmell M.J., Ferguson E.L., Childs S.R., O'Connor M.B.;
 RT "The Drosophila dorsal-ventral patterning gene tolloid is related to
 RL human bone morphogenetic protein 1.";
 RL Cell 67:469-481 (1991).
 RN (2)
 RN SEQUENCE FROM N.A.
 RX MEDLINE=95324373; PubMed=7600963;
 RA Finelli A.U., Bossie C.A., Xie T., Padgett R.W.;
 RT "Mutational analysis of the Drosophila tolloid gene, a human BMP-1
 RL homolog.";
 RL Development 120:861-870 (1994).
 RN (3)
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Cantor A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A.L., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Ewing-Delgado C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 RA Jaitani M., Kalish F., Kapen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K.A., Nussekin D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 CC -1- FUNCTION: Required for normal dorsal development. TLD may interact
 CC physically with Dp-C protein.
 CC -1- MISCELLANEOUS: Mutations in TLD lead to a partial transformation
 CC of dorsal ectoderm into ventral ectoderm.
 CC -1- SIMILARITY: Belongs to peptidase family M12A.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 5 CUB domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M76976; AAA28491.1; -.
 DR EMBL: U04239; AAC64482.1; -.
 DR EMBL: AE003749; AAF56329.1; -.
 DR PIR: A39288; A39288.
 DR HSRP: P00742; 1HCG.
 DR MEROPS: M12.010; -.
 DR Pflam: PF00431; CUB; 5.
 DR Pflam: PF00008; EGF; 2.
 DR PRINTS: PR00480; ASTACIN.
 DR SMART: SM00042; CUB; 5.
 DR SMART: SM00179; EGF CA; 2.
 DR SMART: SM00235; ZDMC; 1.
 DR PROSITE: PS00010; ASX HYDROXYL; 2.
 DR PROSITE: PS01180; CUB; 5.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS00026; EGF_3; 2.
 DR PROSITE: PS01187; EGF CA; 2.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Developmental protein; Hydrolyase; Metalloprotease; Zinc; Glycoprotein;
 KW EGF-like domain; Calcium; Signal; Repeat; Zymogen.
 FT SIGNAL 1 27
 FT PROPEP 28 126
 FT CHAIN 127 1057
 FT
 FT DOMAIN 127 329
 FT DOMAIN 330 467
 FT DOMAIN 468 580
 FT DOMAIN 581 621
 FT DOMAIN 624 742
 FT DOMAIN 743 783
 FT DOMAIN 787 899
 FT DOMAIN 900 1016
 FT DOMAIN 1016 1221
 FT METAL 1222 222
 FT ACT SITE 222 225
 FT METAL 225 231
 FT METAL 231 237
 FT SITE 235 317
 FT SITE 315 317


```

FT DISULFID 330 380 BY SIMILARITY.
FT DISULFID 407 429 BY SIMILARITY.
FT DISULFID 468 495 BY SIMILARITY.
FT DISULFID 522 544 BY SIMILARITY.
FT DISULFID 585 596 BY SIMILARITY.
FT DISULFID 592 605 BY SIMILARITY.
FT DISULFID 607 620 BY SIMILARITY.
FT DISULFID 624 652 BY SIMILARITY.
FT DISULFID 683 706 BY SIMILARITY.
FT DISULFID 747 758 BY SIMILARITY.
FT DISULFID 754 767 BY SIMILARITY.
FT DISULFID 787 813 BY SIMILARITY.
FT DISULFID 840 862 BY SIMILARITY.
FT DISULFID 900 930 BY SIMILARITY.
FT DISULFID 957 979 BY SIMILARITY.
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 533 533 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 781 781 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 854 854 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1057 AA; 120575 MW; 76F4B5AB7996FBA CRC64;

```

```

Query Match 25.2%; Score 217.5; DB 1; Length 1057;
Best Local Similarity 38.4%; Pred. No. 1.7e-14;
Matches 43; Conservative 21; Mismatches 47; Indels 1; Gaps 1;

```

```

Qy 39 VQCGTWTHAGGIFTSPNYSKYPPDECIYITEAARQCIIEYFDEKYSIESWEK 98
Db 466 VVCGGDLKLTQDSINPMYDWPKECVRITAPNHQVAKF-QSFELEKRDGCAY 524
Qy 99 DHIEVDGPGFSPFIIGFCGQNPVIXSSGRFLWIKFADGELSMGFSFA 150
Db 525 DFEVRIDGNHSDSRIGFCGDKLPNIKTRSNQWIRFVDSVQKGFSA 576

```

```

RESULT 11
NRPI RAT STANDARD; PRT; 922 AA.
AC 09QWJ3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neutrophilin-1 precursor (Vascular endothelial cell growth factor 165
  receptor).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=9743085; PubMed=9288754;
RA Kolodkin A.L., Levengood D.V., Rowe E.G., Tai Y.-T., Giger R.D.,
RA Ginty D.D.;
RT "Neutrophilin is a semaphorin III receptor.";
RL Cell 90:753-762(1997).
CC -1- FUNCTION: Receptor involved in the development of the
CC cardiovascular system, in angiogenesis, in the formation of
CC certain neuronal circuits and in organogenesis outside the nervous
CC system. It mediates the chemorepulsive activity of semaphorins. It
CC binds to semaphorin 3A, the PLGF-2 isoform of pGF, the VEGF-165
CC isoform of VEGF and VEGF-8. Coexpression with KDR results in
CC increased VEGF-165 binding to KDR as well as increased chemotaxis.
CC It may regulate VEGF-induced angiogenesis (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Found in the embryonic nervous system.
CC -1- SIMILARITY: Belongs to the neutrophilin family.
CC -1- SIMILARITY: Contains 2 CUB domains.

```

```

CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
CC -1- SIMILARITY: Contains 1 MAM domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

```

DR EMBL; AF016296; AAC3337.1; -.
DR HSSP; P12259; IC2T.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000421; FA58 C.
DR InterPro; IPR008979; Gal Bind like.
DR InterPro; IPR009988; MAM_domain.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR Pfam; PF00629; MAM; 1.
DR PRINTS; PR0020; MAMDOMAIN.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00231; FA58C; 2.
DR SMART; SM0137; MAM; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00022; FA58C_3; 2.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS50060; MAM_2; 1.
DR KAM Receptor.
DR KAM Angiogenesis; Transmembrane; Glycoprotein; Neutrone; Signal; Repeat;
KW SIGNAL.
FT CHAIN 1 21 POTENTIAL.
FT TRANSMEM 22 922 NEUTROFILIN-1.
FT TRANSMEM 856 880 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 881 922 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 141 CUB 1.
FT DOMAIN 147 265 CUB 2.
FT DOMAIN 275 424 F5/8 TYPE C 1.
FT DOMAIN 431 583 F5/8 TYPE C 2.
FT DOMAIN 645 811 MAM.
FT DISULFID 27 54 PROBABLE.
FT DISULFID 82 104 PROBABLE.
FT DISULFID 147 173 PROBABLE.
FT DISULFID 206 228 PROBABLE.
FT DISULFID 275 424 BY SIMILARITY.
FT DISULFID 431 583 BY SIMILARITY.
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 841 841 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 922 AA; 103082 MW; CC6F82AD098B0F2E CRC64;

```

```

Query Match 25.2%; Score 217; DB 1; Length 922;
Best Local Similarity 40.2%; Pred. No. 1.6e-14;
Matches 47; Conservative 20; Mismatches 48; Indels 2; Gaps 2;

```

```

Qy 37 KSVQCGTWTHAGGIFTSPNYSKYPPDECIYITEA-APROCIIEYFDEKYSIESWE 95
Db 23 RSDKCGGTYIKIENPGYITSPGYSHYSPSEKCMWLQAPBPYQRIIMNFPHDLEDR-D 81
Qy 96 CKFDHIEVRDGPFPSPFIIGFCGQNPVIXSSGRFLWIKFADGELSMGFSARY 152
Db 82 CKDYVEVLDGENEGRLWKGKFGKLPSPVSSGPFIFKVDYDYTHGAGFSIRY 138

```

```

RESULT 12
NRPI MOUSE STANDARD; PRT; 923 AA.
AC P97333;
DT 01-NOV-1997 (Rel. 35, Created)

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuropilin-1 precursor (A5 protein).
 GN NRP1 OR NRP.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 OK NCBI_TaxID=10090;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Embryonic brain;
 RX MEDLINE=96353149; PubMed=8748368;
 RA Kanakami A., Kiteukawa T., Takagi S., Fujisawa H.;
 RT "Developmentally regulated expression of a cell surface protein,
 RL neuropilin, in the mouse nervous system.";
 RL J. Neurobiol. 29:1-17(1996).
 CC -1- FUNCTION: Receptor involved in the development of the
 CC cardiovascular system, in angiogenesis, in the formation of
 CC certain neuronal circuits and in organogenesis outside the nervous
 CC system. It mediates the chemorepulsant activity of semaphorins. It
 CC binds to semaphorin 3A, the p165-2 isoform of RGF, the VEGF-165
 CC isoform of VEGF and VEGF-B. Coexpression with KDR results in
 CC increased VEGF-165 binding to KDR as well as increased chemotaxis.
 CC It may regulate VEGF-induced angiogenesis (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Nervous system.
 CC -1- SIMILARITY: Belongs to the neuropilin family.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC -1- SIMILARITY: Contains 2 P5/8 type C domains.
 CC -1- SIMILARITY: Contains 1 MAM domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D50086; BAA08789.1; -;
 DR HSSP; P12259; 1CZT.
 DR MGD; MGI:106206; Nrp.
 DR GO; GO:0017154; F:semaphorin receptor activity; IGI.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR000421; FAS8_C.
 DR InterPro; IPR008979; Gal_Bind_like.
 DR InterPro; IPR000998; MAM_domain.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00754; F5_F8_type_C; 2.
 DR Pfam; PF00629; MAM; 1.
 DR PRINTS; PR00020; MAMDOMAIN.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00231; FAS8C; 2.
 DR SMART; SM00137; MAM; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS01285; FAS8C_1; 2.
 DR PROSITE; PS01286; FAS8C_2; 2.
 DR PROSITE; PS50022; FAS8C_3; 2.
 DR PROSITE; PS00740; MAM_1; 1.
 DR PROSITE; PS50060; MAM_2; 1.
 DR PROSITE; PS50060; MAM_2; 1.
 DR Angiogenesis; Transmembrane; Glycoprotein; Neurone; Signal; Repeat;
 KW Receptor.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 923 NEUROPIILIN-1.
 FT DOMAIN 22 856 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 857 879 POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 880 923
 FT DOMAIN 27 141 CUB 1.
 FT DOMAIN 147 265 CUB 2.
 FT DOMAIN 275 424 F5/8 TYPE C 1.
 FT DOMAIN 431 583 F5/8 TYPE C 2.
 FT DOMAIN 645 811 MAM.
 FT DISULFID 27 54 PROBABLE.

FT DISULFID 82 104 PROBABLE.
 FT DISULFID 147 173 PROBABLE.
 FT DISULFID 206 228 PROBABLE.
 FT DISULFID 275 424 BY SIMILARITY.
 FT DISULFID 431 583 BY SIMILARITY.
 FT DISULFID 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 923 AA; 103020 MW; 06448BA170796808 CRC64;
 Query Match 25.1%; Score 216; DB 1; Length 923;
 Best Local Similarity 40.2%; Pred. No. 2.1e-14;
 Matches 47; Conservative 20; Mismatches 48; Indels 2; Gaps 2;
 QY 37 KSVGCGTMTAEGGIFTPSPVPSKYPDRDREGIYIEA-APRQCTELVDEKYSIEPSWE 95
 DB 23 RSDKGGTGTIKENPEYVLSFGPHSHSEKCEWLIQAPYQRIINPNPHFDEDR-D 81
 QY 96 CKFDHIEVRDGPFGPSPITIGRCQGNPVIKSGRPIMIKFPADGELSMGFSARY 152
 DB 82 CKYDVEYIDENEGRIMKGRCKRIAPSPVSSGPFPIKFKVSDVETHAGFSIRY 138
 RESULT 13
 BMMH STRPU STANDARD; PRT; 639 AA.
 ID BMMH STRPU
 AC P98069;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2001 (Rel. 40, Last annotation update)
 DE Bone morphogenetic protein 1 homolog precursor (EC 3.4.24.-) (SUBMP).
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoida; Echinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OC NCBI_TaxID=7668;
 RN NCBI_TaxID=7668;
 RP
 RC SEQUENCE FROM N.A.
 RX MEDLINE=94215496; PubMed=8162855;
 RA Hwang S.-P., Partin J.S., Lemariz W.J.;
 RT "Characterization of a homolog of human bone morphogenetic protein 1
 RT in the embryo of the sea urchin, Strongylocentrotus purpuratus.";
 RT Development 120:559-568(1994).
 CC -1- TISSUE SPECIFICITY: Ectodermal and primary mesenchym cells in
 CC hatched blastula.
 CC -1- DEVELOPMENTAL STAGE: Embryo; highest level before spiculogenesis.
 CC -1- SIMILARITY: Belongs to peptidase family M12A.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L23838; AAA30081.1; -;
 DR HSSP; P00736; 1APQ.
 DR MEROPS; M12.005; -;
 DR InterPro; IPR000152; Asx_hydroxyl_S.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006025; Pept_M_zn_BS.
 DR InterPro; IPR006026; Peptidase_M.
 DR InterPro; IPR001506; Peptidase_M12A.
 DR Pfam; PF01400; Ascatin_1.
 DR Pfam; PF00431; CUB; 2.

DR Pfam: PF00008; EGF 1.
DR PRINTS: PR00480; ASTACTIN.
DR SMART: SM00042; CUB; 2.
DR SMART: SM00179; EGF_CA; 1.
DR SMART: SM00235; ZMG; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS01180; CUB; 2.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF 2; 1.
DR PROSITE: PS50026; EGF 3; 1.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Developmental protein; Repeat; Hydrolase; Protease; Zinc; Embryo;
Metalloprotease; EGF-like domain; Calcium; Signal.
FT SIGNAL 1 23
FT PROPEP 24 109
FT CHAIN 110 639
FT DOMAIN 110 306
FT DOMAIN 307 419
FT DOMAIN 420 531
FT DOMAIN 532 573
FT METAL 197 197
FT ACT SITE 198 198
FT METAL 201 201
FT METAL 207 207
FT DISULFID 536 548
FT DISULFID 544 557
FT DISULFID 559 572
FT CARBOHYD 122 122
FT CARBOHYD 140 140
FT CARBOHYD 317 317
FT CARBOHYD 455 455
SQ SEQUENCE 639 AA; 71893 MW; 59307B265B7894AD CRC64;

Query Match 24 9%; Score 215; DB 1; Length 639;
Best Local Similarity 39.3%; Pred. No. 1.7e-14;
Matches 44; Conservative 19; Mismatches 47; Indels 2; Gaps 2

QY 41 CGTWTKEAGSGIFLTPSPYSPKSPDPDECIYIADPROCIETLYFDKYSITSPSWCKPQH 100
DB 420 CGGHIER-EGSTIGSPHYPPDHPHRSKGCWLLTLPANYIGLSF-QSFLERHETCIYDY 477
QY 101 IEVVDGFGFGSPITIGRFQGGQNPVIVKSGRFLWIKFFADGELSSMGFSARY 152
DB 478 VEVDGDGHEHDSPLIGRYCGYFIPDDIKSTGNKMMVTFVSDGSYNGKGFSADE 529

RESULT 14
FCOL_HUMAN STANDARD; PRT; 449 AA.
ID FCOL_HUMAN
AC Q15113; O14550;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Procollagen C-proteinase enhancer protein precursor (PCPE) (Type I
DE Procollagen COOH-terminal proteinase enhancer) (Type I procollagen C-
DE proteinase enhancer protein).
GN PCOLCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta.
RX MEDLINE=95014462; PubMed=7523404;
RA Takahara K., Kessler E., Binnamov L., Brusel M., Eddy R.L.,
RA Jani-Salt S., Shows T.B., Greenspan D.S.;
RT "Type I procollagen COOH-terminal proteinase enhancer protein:
RT identification, primary structure, and chromosomal localization of the
RT cognate human gene (PCOLCE)".
RU J. Biol. Chem. 269:26280-26285 (1994).
RN [2]

RP REVISIONS TO 56; 154 AND 373.
RA Kessler E.;
RL Unpublished observations (FEB-2000).
[3]
RP SEQUENCE FROM N.A.
RA Tissue-Heart;
RA Hatahara I., Syoufuda K., Harada K., Tomita M., Urakami K., Terai H.,
RA Morisaki N., Saito Y.;
RT "Smooth muscle cell derived procollagen C-protease enhancer protein.";
RL Cell Struct. Funct. 21:662-662(1996).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=99018116; PubMed=9799793;
RA Gloeckner G., Scherer S., Schattevoy R., Boright A.P., Weber J.,
RA Tsui L.-C., Rosenthal A.;
RT "Large-scale sequencing of two regions in human chromosome 7q22:
RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT reveals 17 genes.";
RL Genome Res. 8:1060-1073(1998).
[5]
RP SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RC MEDLINE=99134301; PubMed=9933570;
RX Scott I.C., Clark T.G., Takahara K., Hoffman G.G., Greenspan D.S.;
RT "Structural organization and expression patterns of the human and
RT mouse genes for the type I procollagen COOH-terminal proteinase
RT enhancer protein.";
RL Genomics 55:229-234(1999).
[6]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Collins E.A., Grouse L.H., Derge J.G.,
RA Klausberg R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepleton M., Soares M.B., Donald M.F., Casavant T.P., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lofgren N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gamaralle P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Bakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[7]
RP PARTIAL SEQUENCE, AND CHARACTERIZATION OF INHIBITORY ACTIVITY.
RX MEDLINE=20092917; PubMed=10625689;
RA Mott J.D., Thomas C.L., Rosenbach M.T., Takahara K., Greenspan D.S.,
RA Banda M.J.;
RT "Post-translational proteolytic processing of procollagen C-terminal
RT proteinase enhancer releases a metalloproteinase inhibitor.";
RL J. Biol. Chem. 275:1384-1390(2000).
[8]
RP FUNCTION: Binds to the COOH-terminal propeptide of type I
CC procollagen and enhances procollagen C-proteinase activity.
CC FUNCTION: C-terminal processed part of PCPE (CT-PCPE) may have an
CC metalloproteinase inhibitory activity.
CC SUBCELLULAR LOCATION: Secreted.
CC -1- PMW: C-terminally processed at multiple positions.
CC -1- SIMILARITY: Contains 2 CUB domains.
CC -1- SIMILARITY: Contains 1 NTR domain.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

This Page Blank (usptol)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 18:21:16 ; Search time 21 Seconds

(Without alignments)
723.727 Million cell updates/sec

Title: US-09-887-593A-2

Perfect score: 862

Sequence: 1 MIHGRSVLHVLAFLIILHL.....ADGELSMGFSARVNFPPK 158

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR 78:*
1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	818	94.9	533	JC7985
2	260	30.2	321	T33161
3	227.5	26.4	3623	T08618
4	227.5	26.4	3623	T09456
5	227	26.3	707	JC2218
6	224	26.0	991	I49540
7	220.5	25.6	986	B58788
8	217.5	25.2	1057	A35288
9	216.5	25.1	1464	S58984
10	211	24.5	730	BMH1
11	211	24.5	823	A58788
12	199.5	23.1	1070	T31069
13	199	23.1	579	JC7629
14	198.5	23.0	927	JC0948
15	197.5	22.9	449	A55362
16	193.5	22.4	402	JH0403
17	187.5	21.8	1594	T30549
18	178.5	20.7	699	I54763
19	172.5	20.0	2403	A53886
20	171.5	19.9	1524	T30337
21	170.5	19.8	597	T31352
22	164.5	19.1	2083	T42721
23	162.5	18.9	504	S56745
24	161.5	18.7	3871	T22812
25	152.5	17.7	705	CHH08
26	151	17.5	275	JC6506
27	150	17.4	310	A41735
28	150	17.4	310	S41055
29	149	17.3	1290	A57190

ALIGNMENTS

30	148	17.2	276	2	A47290	TS6-6 homolog B94
31	146	16.9	412	2	S72579	hypothetical prote
32	145	16.8	421	2	T25383	hypothetical prote
33	143.5	16.6	288	2	T33224	hypothetical prote
34	143.5	16.6	417	2	T20199	hypothetical prote
35	142.5	16.5	319	2	I51569	IWS.2 protein - Af
36	142.5	16.5	767	2	JC7018	hypothetical prote
37	141.5	16.4	855	2	T30731	membrane-bound arg
38	137	15.9	1004	2	T30338	oviductin (BC 3.4.
39	131	15.2	416	2	T20448	hypothetical prote
40	128	14.8	383	2	T21946	hypothetical prote
41	125	14.5	770	2	T00203	IDL receptor relat
42	123	14.3	770	2	T00204	IDL receptor relat
43	122.5	14.2	317	2	JC7592	spinal cord-derive
44	121.5	14.1	470	2	T29864	hypothetical prote
45	118	13.7	370	2	JC7591	spinal cord-derive

RESULT 1

JC7985

brain-specific CUB and LDL domain-containing transmembrane protein, BTCL1 protein - M

C/Species: Mus musculus (house mouse)

C/Date: 10-Nov-2003 #sequence_revision 10-Nov-2003 #text_change 24-Nov-2003

C/Accession: JC7985; PC7228

R/Michishita, M.; Ikeda, T.; Nakashiba, T.; Ogawa, M.; Tashiro, K.; Honjo, T.; Doi, K.

Biochem. Biophys. Res. Commun. 306, 680-686, 2003

A/Title: A novel gene, Btcl1, encoding CUB and LDL domains is expressed in restricted

A/Reference number: JC7985; PMID:1281072

A/Accession: JC7985

A/Molecule type: mRNA

A/Residues: 1-533 <MIC>

A/Cross-references: GB:AY138990

A/Experimental source: Cerebellum, 3-week-old

A/Accession: PC7228

A/Molecule type: protein

A/Residues: 1-61 <MI2>

C/Comment: This protein belongs to a new class of brain-specific type I transmembrane p

C/Genetics:

A/Keywords: BTCL1 protein; CUB domain; LDL domain; transmembrane protein

Query Match

Best Local Similarity 94.2%; Pred. No. 6, 1e-76;

Matches 147; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY	1	MIHGRSVLHVLAFLIILHSGATKKGTQTSFQSVQCGTWTKEAGGIFSPNPS	60
DB	1	MIYGRSLPHILASLIILHSSGATKKGTQTSFQSVQCGTWTKEAGGIFSPNPS	60
QY	61	KYPPRECVYIIIEAAROCIEIYFDEKYSIEPSWECKFDHIEVRDGPFGFSPIIRFCQ	120
DB	61	KYPPRECVYIIIEAAROCIEIYFDEKYSIEPSWECKFDHIEVRDGPFGFSPIIRFCQ	120
QY	121	QNPVYIKSGRFLWIKFPADELSMGFSARVNFPP	156
DB	121	QNPVYIKSGRFLWIKFPADELSMGFSARVNFPP	156

RESULT 2

T33161

hypothetical protein K035.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C/Accession: T33161

R/Becker, M.; Graves, T.; Fronick, B.

submitted to the EMBL Data Library, May 1998

A/Description: The sequence of C. elegans coamid K035.

A/Reference number: Z21294

A/Accession: T33161

A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-321 <BEC>
 A:Cross-references: EMBL:AF067619; PIDN:AACT17565.1; GSPDB:GN00019; CESP:K03E5.1
 A:Experimental source: strain Bristol N2; clone K03E5
 C:Genetics:
 A:Gene: CESP:K03E5.1
 A:Map position: 1
 A:Introns: 25/1; 47/1; 84/2; 112/3; 131/3; 162/1; 197/1; 275/2

Query Match 30.2%; Score 260; DB 2; Length 321;
 Best Local Similarity 43.5%; Pred. No. 6,3e-19;
 Matches 50; Conservative 20; Mismatches 37; Indels 8; Gaps 2;

QY 52 IFTSPNPKYPPDRECIYIIEAAROCIELYDEKYSIEPSM-----ECKFDHIEVR 104
 Db 75 IFTSPNPKYPPDRECIYIIEAAROCIELYDEKYSIEPSM-----ECKFDHIEVR 134

QY 105 DEFGFSPITGRGCGQGNP-VIKSGRFLMIKFPADGELSMGFSARYNFTGK 158
 Db 135 DGRYGFSPITGRGCGQGNP-VIKSGRFLMIKFPADGELSMGFSARYNFTGK 189

RESULT 3
 T08618
 Intrinsic factor-B12 receptor CUBILIN precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 02-Aug-2002
 C:Accession: T08618
 R:Moestrup, S.K.; Kozzyraki, R.; Kristiansen, M.; Kayser, J.H.; Rasmussen, H.H.; Brault, J.; Biol. Chem. 273, 5235-5242, 1998
 A:Title: The intrinsic factor-vitamin B12 receptor and target of teratogenic antibodies
 A:Reference number: Z16459; MUID:98148073; PMID:9478979
 A:Accession: T08618
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-3623 <MOE>
 A:Cross-references: EMBL:AF022247; NID:G3834379; PIDN:AACT1661.1; PID:G3834380
 C:Genetics:
 A:Gene: CUBILIN
 C:Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology
 C:Keywords: egg yolk; endocytosis; glycoprotein; intestine; kidney; peripheral membrane
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-3623/Product: intrinsic factor-B12 receptor CUBILIN #status predicted <MAT>
 F:133-164/Domain: EGF homology <EGF1>
 F:436-467/Domain: EGF homology <EGF>

Query Match 26.4%; Score 227.5; DB 2; Length 3623;
 Best Local Similarity 37.5%; Pred. No. 2.1e-14;
 Matches 45; Conservative 26; Mismatches 46; Indels 3; Gaps 2;

QY 37 KSVQCG--TWTKHAE--GIFTSPNPKYPPDRECIYIIEAAROCIELYDEKYSIEPS 93
 Db 2213 KSIACGGNVYIHDAADSDGYLSPNPKYPPDRECIYIIEAAROCIELYDEKYSIEPS 2272

QY 94 WECKFDHIEVRDGPFGFSPITGRGCGQGNPVIKSGRFLMIKFPADGELSMGFSARYN 153
 Db 2273 PNCSTNVLELRDGVSDAPILSKFCGTLPSSQWSSGVWYLRFRSDNSPTHWGFKAKYS 2332

RESULT 4
 T09456
 Intrinsic factor-B12 receptor Cubilin precursor - human
 C:Species: Homo sapiens (man)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Aug-2002
 C:Accession: T09456
 R:Kozzyraki, R.; Kristiansen, M.; Sliacharoglu, A.; Hansen, C.; Jacobsen, C.; Tommerup, N.
 Blood 91, 3593-3600, 1998
 A:Title: The human intrinsic factor-vitamin B12 receptor, cubilin: Molecular characterization.
 A:Reference number: Z16677; MUID:98241400; PMID:9572993
 A:Accession: T09456
 A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-3623 <KOZ>
 A:Cross-references: EMBL:AF034611; NID:G3929528; PIDN:AAAC82612.1; PID:G3929529
 C:Genetics:
 A:Map position: 10p12
 C:Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology
 C:Keywords: receptor; vitamin B12 uptake
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-3623/Product: intrinsic factor-B12 receptor #status predicted <MAT>
 F:436-467/Domain: EGF homology <EGF>

Query Match 26.4%; Score 227.5; DB 2; Length 3623;
 Best Local Similarity 35.1%; Pred. No. 2.1e-14;
 Matches 46; Conservative 29; Mismatches 47; Indels 9; Gaps 3;

QY 37 KSVQCG--TWTKHAE--GIFTSPNPKYPPDRECIYIIEAAROCIELYDEKYSIEPS 93
 Db 2213 KSIACGGNVYIHDAADSDGYLSPNPKYPPDRECIYIIEAAROCIELYDEKYSIEPS 2272

QY 94 WECKFDHIEVRDGPFGFSPITGRGCGQGNPVIKSGRFLMIKFPADGELSMGFSARYN 153
 Db 2273 PNCSTNVLELRDGVSDAPILSKFCGTLPSSQWSSGVWYLRFRSDNSPTHWGFKAKYS 2332

QY 154 F-----TPGK 158
 Db 2333 IAQCGGRVPCQ 2343

RESULT 5
 JC2218
 Procollagen C-endopeptidase (EC 3.4.24.19) - African clawed frog (clone 22AN)
 N:Alternate names: bone morphogenic protein 1
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Sep-1999
 C:Accession: JC2218
 R:Maeno, M.; Xue, Y.; Wood, T.I.; Ong, R.C.; Kung, H.
 Gene 134, 257-261, 1993
 A:Title: Cloning and expression of cDNA encoding Xenopus laevis bone morphogenetic prot.
 A:Reference number: JC2218; MUID:94085787; PMID:8262384
 A:Accession: JC2218
 A:Molecule type: mRNA
 A:Residues: 1-707 <MAB>
 A:Cross-references: GB:L1249; NID:G406540; PIDN:AAAL6313.1; PID:G406541
 C:Comment: This protein induces ectopic cartilage formation in vivo.
 C:Superfamily: procollagen C-endopeptidase; astacin homology; C1r/C1s repeat homology;
 C:Keywords: beta-hydroxyasparagine; glycoprotein; hydrolase; metalloproteinase; zinc
 F:93-284/Domain: astacin homology <AST>
 F:285-397/Region: complement 1r/1s-like repeat
 F:398-510/Region: complement 1r/1s-like repeat
 F:398-510/Region: complement 1r/1s-like repeat
 F:514-550/Domain: C1r/C1s repeat homology <C1R2>
 F:554-666/Region: complement 1r/1s-like repeat
 F:554-666/Region: complement 1r/1s-like repeat
 F:554-666/Domain: C1r/C1s repeat homology <C1R3>
 F:662-105,295,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:176,180,186,235/Binding site: zinc (His, His, Tyr) #status predicted
 F:177/Active site: Glu #status predicted
 F:528/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 26.3%; Score 227; DB 2; Length 707;
 Best Local Similarity 40.2%; Pred. No. 3.7e-15;
 Matches 45; Conservative 18; Mismatches 47; Indels 2; Gaps 2;

QY 41 CGTWTKHAE--GIFTSPNPKYPPDRECIYIIEAAROCIELYDEKYSIEPSWECKFDH 100
 Db 398 CGGEVKKYSGH--QSPNPKYPPDRECIYIIEAAROCIELYDEKYSIEPSWECKFDH 455

QY 101 IEVRDGPFGFSPITGRGCGQGNPVIKSGRFLMIKFPADGELSMGFSARYN 152
 Db 456 IEVRDGPFGFSPITGRGCGQGNPVIKSGRFLMIKFPADGELSMGFSARYN 507

RESULT 6

149540
 procollagen C-endopeptidase (EC 3.4.24.19) precursor, tolloid-like splice form - mouse
 C.Species: Mus musculus (house mouse)
 C.Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
 C.Accession: I49540
 R.Fukagawa, M.; Noboru, S.; Hogan, B.L.M.; Jones, C.M.
 Dev. Biol. 163, 175-183, 1994
 A>Title: Embryonic expression of mouse bone morphogenetic protein-1 (BMP-1) which is related to human BMP-1
 A.Reference number: I49540; MUID:94229342; PMID:8174772
 A.Accession: I49540
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: mRNA
 A.Residues: 1-991 <RES>
 A.Cross-references: GB:I24755; NID:9439606; PIDN:AAA37306.1; PID:9439607
 C.Genetics:
 A.Gene: Bmp-1
 C.Superfamily: procollagen C-endopeptidase; astacin homology; Clr/Cls repeat homology; F115-326/Domain: astacin homology <AST>
 C.Keywords: hydrolase; metalloproteinase; zinc
 F:115-326/Domain: astacin homology <AST>
 F:556-592/Domain: EGF homology <EG1>
 F:596-705/Domain: Clr/Cls repeat homology <CLR>
 F:712-747/Domain: EGF homology <EG2>
 F:218-222,228,277/Binding site: zinc (His, His, His, Tyr) #status predicted
 F:219/Active site: Glu #status predicted

Query Match 26.0%; Score 224; DB 2; Length 991;
 Best Local Similarity 33.6%; Pred. No. 1,1e-14;
 Matches 43; Conservative 27; Mismatches 42; Indels 16; Gaps 2;

QY 26 GTEKOTSTQKSVGCGTWTKHAEAGIFTSPPVPSKYPPDRRCITITIAAPQCIELYPD 85
 DB 751 GGEHKVTS-----GTTSPMMDKYPKSKCTWALSTPGHRYVLTIV 795

QY 86 EKYSIEPSWECKEHLIEVDGFGFSPILGRFCGQNPVIVSSGRFLWIKFPADGELES 145
 DB 796 E-MDIESQECAYDHLIEVDGDAKAPVLRFCGSKKPEPVATAGRMFLRYSDNSVQR 854

QY 146 MGFSAHYN 153
 DB 855 KGFOASHS 862

RESULT 7
 B58788
 procollagen C-endopeptidase (EC 3.4.24.19) precursor, tolloid-like splice form - human
 N.Alternate names: bone morphogenic protein 1, tolloid-like splice form
 C.Species: Homo sapiens (man)
 C.Date: 28-Mar-1998 #sequence_revision 09-Apr-1998 #text_change 16-Jul-1999
 C.Accession: A37278; B58788
 R.Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitscock, L.M.; Whitters, M.J.; Kriz, R.W.; Hew Science 242, 1528-1534, 1988
 A>Title: Novel regulators of bone formation: molecular clones and activities.
 A.Reference number: A37278; MUID:89072730; PMID:3201241
 A.Accession: A37278
 A.Molecule type: mRNA
 A.Residues: 1-702, EKRPALPPRRPHQLKFRVQKRNRTPO' <W02>
 A.Cross-references: GB:M22488; NID:9179499; PIDN:AAA51833.1; PID:9179500
 R.Takahara, K.; Lyons, G.E.; Greenspan, D.S.
 J. Biol. Chem. 269, 32572-32578, 1994
 A>Title: Bone morphogenetic protein-1 and a mammalian tolloid homologue (MTLD) are encoded by the same gene.
 A.Reference number: A58788; MUID:95096114; PMID:7798260
 A.Accession: B58788
 A.Molecule type: mRNA
 A.Residues: 703-986 <TAK>
 A.Cross-references: GB:I35279; NID:9619860; PIDN:AA41710.1; PID:9619861
 C.Genetics:
 A.Gene: GDB: BMP-1
 A.Cross-references: GDB:125203; OMIM:112264
 A.Map position: 8p21-8p21
 C.Function:
 A.Description: catalyzes hydrolysis of the carboxyl-terminal propeptide of collagen type I
 C.Superfamily: procollagen C-endopeptidase; astacin homology; Clr/Cls repeat homology; F115-326/Domain: astacin homology <AST>
 C.Keywords: alternative splicing; beta-hydroxyasparagine; bone; calcium; duplication; gl

F:1-22/Domain: signal sequence #status predicted <SIG>
 F:33-986/Product: procollagen C-endopeptidase tolloid-like splice form #status predicted
 F:730-321/Domain: astacin homology <AST>
 F:322-431/Domain: Clr/Cls repeat homology <CLR1>
 F:435-544/Domain: Clr/Cls repeat homology <CLR2>
 F:551-587/Domain: EGF homology <EG1>
 F:591-700/Domain: Clr/Cls repeat homology <CLR3>
 F:707-742/Domain: EGF homology <EG2>
 F:747-856/Domain: Clr/Cls repeat homology <CLR4>
 F:860-973/Domain: Clr/Cls repeat homology <CLR5>
 F:911-142,332,363,599/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F:163-319,185-205,322-348,375-397,435-461,488-510,551-563,569-572,574-587,591-617,644-660/Binding site: zinc (His, His, His, Tyr) #status predicted
 F:213,217,223,272/Binding site: zinc (His, His, His, Tyr) #status predicted
 F:214/Active site: Glu #status predicted
 F:565,720/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 25.6%; Score 220.5; DB 1; Length 986;
 Best Local Similarity 38.2%; Pred. No. 2.5e-14;
 Matches 39; Conservative 24; Mismatches 38; Indels 1; Gaps 1;

QY 51 GFTSPNPSKYPDPRECIYITIAAPQCIEYFDEKYSIEPSWECKPHIVRDGPGFR 110
 DB 756 GTTSPNMDKYPKSKCTWALSTPGHRYVLTIV 814

QY 111 SPITGRFCGQNPVIVSSGRFLWIKFPADGELESMSGFARY 152
 DB 815 AEVLGRFCGSKKPEPVATAGRMFLRYSDNSVQRKGFASH 856

RESULT 8
 A39288
 dorsal-ventral patterning protein tolloid (EC 3.4.24.-) - fruit fly (Drosophila melanogaster)
 C.Species: Drosophila melanogaster
 C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C.Accession: A39288
 R.Shmell, M.J.; Ferguson, E.L.; Childs, S.R.; O'Connor, M.B.
 Cell 67, 469-481, 1991
 A>Title: The Drosophila dorsal-ventral patterning gene tolloid is related to human bone morphogenetic protein-1.
 A.Reference number: A39288; MUID:92034970; PMID:1840509
 A.Accession: A39288
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-1057 <SHI>
 A.Cross-references: GB:M6976; NID:9157305; PIDN:AAA28491.1; PID:9157306
 C.Genetics:
 A.Gene: FlyBase:tlid
 A.Cross-references: FlyBase:FBgn0003719
 C.Superfamily: dorsal-ventral patterning protein tolloid; astacin homology; Clr/Cls repeat homology <CLR>
 C.Keywords: duplication; hydrolase; metalloproteinase; zinc
 F:136-329/Domain: astacin homology <AST>
 F:352-464/Domain: Clr/Cls repeat homology <CLR1>
 F:468-578/Domain: Clr/Cls repeat homology <CLR2>
 F:585-620/Domain: EGF homology <EG1>
 F:624-740/Domain: Clr/Cls repeat homology <CLR3>
 F:747-782/Domain: EGF homology <EG2>
 F:787-896/Domain: Clr/Cls repeat homology <CLR4>
 F:900-1013/Domain: Clr/Cls repeat homology <CLR5>
 F:121,225,231,280/Binding site: zinc (His, His, His, Tyr) #status predicted
 F:222/Active site: Glu #status predicted

Query Match 25.2%; Score 217.5; DB 1; Length 1057;
 Best Local Similarity 38.4%; Pred. No. 5.6e-14;
 Matches 43; Conservative 21; Mismatches 47; Indels 1; Gaps 1;

QY 39 VOCGWTAKAEGGIFTSPPVPSKYPDPRECIYITIAAPQCIEYFDEKYSIEPSWECKR 98
 DB 466 VVCGGDLKTKQSDISDPYPMDMPKCECVRITAPDNHQAALKF-GSELEKHDGCAY 524

QY 99 DHIEVRDGFPGFSPITIGRFGQNPVIVSSGRFLWIKFPADGELESMSGFS 150
 DB 525 DVEIRDGNHSDRLIGRFGCKLPNPKITRNSQVYIRVSDSSVQKGFSA 576


```

Db          256 GR 257
:
:
RESULT 14
J00948
A5 antigen precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JH0466; J00948
R:Takagi, S.; Hirata, T.; Agata, K.; Mochii, M.; Eguchi, G.; Fujisawa, H.
Neuron 7, 295-307, 1991
A:Title: The A5 antigen, a candidate for the neuronal recognition molecule, has homology
A:Reference number: JH0466; MUID:91337458; PMID:1908252
A:Accession: JH0466
A:Molecule type: mRNA
A:Residues: 1-927 <TRK>
A:Cross-references: GB:D10467; GB:D01077; NID:g222962; PIDN:BA01260.1; PID:g222963
A:Experimental source: tadpole, brain
A:Note: this protein has motifs homologous to complement components C1r and C1s and to
C:Comment: This protein is a neuronal cell surface molecule involved in the neuronal re-
C:Superfamily: Xenopus A5 antigen, C1r/C1s repeat homology; discoidin I amino-terminal
C:Keywords: duplication; glycoprotein; C1r/C1s repeat homology; discoidin I amino-terminal
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-927/Product: A5 antigen #status predicted <ASA>
F:27-138/Domain: C1r/C1s repeat homology <C1R1>
F:147-262/Domain: C1r/C1s repeat homology <C1R2>
F:274-424/Domain: discoidin I amino-terminal homology <DN1>
F:430-564/Domain: discoidin I amino-terminal homology <DN2>
F:646-812/Domain: MAM homology <MAM>
F:861-883/Domain: transmembrane #status predicted <TM>
F:150,261,300,523,844/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      23.0%; Score 198.5; DB 1; Length 927;
Best Local Similarity 32.4%; Pred. No. 4.3e-12;
Matches 44; Conservative 26; Mismatches 53; Indels 13; Gaps 3;

QY 32 TSETQKQVCGGTWTKHAEGLFSTSNPKPPRECIYITEAAR-QCIETLPDEKSI 90
Db 18 SSMASRNDKCGDITIKTSPSYLTSGYPHSTPSQRCWMLQAEVHYQRIIMNFPHPDL 77
QY 91 EPSWCEKFDHIEVADGFGFSPITIGRFGQGNPPIVSKSGRFLWIKFPADGELBSMGPSA 150
Db 78 EDR-ECKYDVVEVIDGNNANGQLGKXCCKRIAPSPVSTGSIIRFVSVDYETGAGRSI 136
QY 151 RY-----NFT 155
Db 137 RYEVFKTGPECSRNF 152

RESULT 15
A55362
procollagen I C-proteinase enhancer protein precursor - human
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 05-Nov-1999
C:Accession: A55362
R:Takahara, K.; Kessler, E.; Biniaminov, L.; Brusel, M.; Eddy, R.L.; Jani-Sait, S.; Sho
U. J. Biol. Chem. 269, 26280-26285, 1994
A:Title: Type I procollagen COOH-terminal proteinase enhancer protein: identification, i
A:Reference number: A55362; MUID:95014462; PMID:7523404
A:Accession: A55362
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-449 <TRK>
A:Cross-references: GB:I33799; NID:g642307; PIDN:AAA61949.1; PID:g642308
C:Genetics:
A:Gene: GDB:PCOLCE
A:Cross-references: GDB:305468; OMIM:600270
A:Map position: 7q21.3-7q22
C:Superfamily: C1r/C1s repeat homology
C:Keywords: extracellular protein; glycoprotein; pyroglutamic acid
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-449/Product: #status predicted <MAT>

```

F:37-146/Domain: C1r/C1s repeat homology <C1R1>
F:159-270/Domain: C1r/C1s repeat homology <C1R2>
F:26/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:29,431/Banding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	22.9%;	Score 197.5;	DB 2;	Length 449;
Post Total Similarity	36.4%;	Pred No 3	4e-12;	

```

QY 16 ILHLGAKKKGEKQOTSETLOKSVQCGMWTGHAEGGIPTSPVP-SKYPDRECIYIEA 74
Db 143 LLMVSGRATSSSEHFP-----CGRLGAAGQTL--TTPWPSDSYDPGISCWHITA 192
QY 75 APROCIETLYPDBKYSIEPSMECKFPHIEVRDGPFG-FSPIIGRFQCGOONPYIKSSGRFL 133
Db 193 PPDVIALITF-EKFLPEEDPYCRYISVSVYFNCVAVSDBDRRLCKFCGDVAPGSISSGNEL 251
QY 134 WIKFPADGELBSMGFSARYNPTP 156
Db 252 LVQFVSDLSVTNADGFSASGYKTLTP 274

```

Search completed: May 18, 2004, 18:24:52
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2004, 18:23:47 ; Search time 43 Seconds

(without alignments)
1022,449 Million cell updates/sec

Title: US-09-887-593a-2

Perfect score: 862

Sequence: 1 MIHRSVLIHVASLIILHLIS.....ADGHEISMGFSARVPTPGK 158

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	862	100.0	158	9 US-09-887-593-2	Sequence 2, Appl
2	851	98.7	158	15 US-10-138-588-38	Sequence 38, Appl
3	851	98.7	533	13 US-10-041-406-2	Sequence 27, Appl
4	851	98.7	533	14 US-10-162-435-27	Sequence 27, Appl
5	479	55.6	130	9 US-09-764-853-653	Sequence 653, Appl
6	479	55.6	141	10 US-09-764-881-116	Sequence 116, Appl
7	479	55.6	141	10 US-09-764-881-116	Sequence 116, Appl
8	479	55.6	141	15 US-10-242-747-116	Sequence 116, Appl
9	470	54.5	352	12 US-10-276-774-2238	Sequence 2238, Ap
10	470	54.5	525	12 US-10-147-423-114	Sequence 114, Appl
11	470	54.5	525	12 US-10-145-127-114	Sequence 114, Appl
12	470	54.5	525	12 US-10-160-503-114	Sequence 114, Appl
13	470	54.5	525	12 US-10-143-118-114	Sequence 114, Appl
14	470	54.5	525	12 US-10-144-993-114	Sequence 114, Appl
15	470	54.5	525	12 US-10-158-787-114	Sequence 114, Appl

16	470	54.5	525	12	US-10-140-024-114	Sequence 114, Appl
17	470	54.5	525	12	US-10-140-808-114	Sequence 114, Appl
18	470	54.5	525	12	US-10-152-405-114	Sequence 114, Appl
19	470	54.5	525	12	US-10-127-852A-114	Sequence 114, Appl
20	470	54.5	525	12	US-10-127-900A-114	Sequence 114, Appl
21	470	54.5	525	12	US-10-128-685A-114	Sequence 114, Appl
22	470	54.5	525	12	US-10-131-820A-114	Sequence 114, Appl
23	470	54.5	525	12	US-10-142-886-114	Sequence 114, Appl
24	470	54.5	525	12	US-10-146-728-114	Sequence 114, Appl
25	470	54.5	525	12	US-10-146-768-114	Sequence 114, Appl
26	470	54.5	525	12	US-10-147-499-114	Sequence 114, Appl
27	470	54.5	525	12	US-10-157-798-114	Sequence 114, Appl
28	470	54.5	525	12	US-10-028-072-114	Sequence 114, Appl
29	470	54.5	525	14	US-10-121-049-114	Sequence 114, Appl
30	470	54.5	525	14	US-10-123-904-114	Sequence 114, Appl
31	470	54.5	525	14	US-10-140-470-114	Sequence 114, Appl
32	470	54.5	525	14	US-10-175-746-114	Sequence 114, Appl
33	470	54.5	525	14	US-10-176-918-114	Sequence 114, Appl
34	470	54.5	525	14	US-10-176-921-114	Sequence 114, Appl
35	470	54.5	525	14	US-10-137-865-114	Sequence 114, Appl
36	470	54.5	525	14	US-10-140-474-114	Sequence 114, Appl
37	470	54.5	525	14	US-10-142-431-114	Sequence 114, Appl
38	470	54.5	525	14	US-10-143-114-114	Sequence 114, Appl
39	470	54.5	525	14	US-10-140-002-114	Sequence 114, Appl
40	470	54.5	525	14	US-10-142-419-114	Sequence 114, Appl
41	470	54.5	525	14	US-10-123-262-114	Sequence 114, Appl
42	470	54.5	525	14	US-10-142-423-114	Sequence 114, Appl
43	470	54.5	525	14	US-10-213-181-12	Sequence 12, Appl
44	470	54.5	525	14	US-10-121-050-114	Sequence 114, Appl
45	470	54.5	525	14	US-10-141-755-114	Sequence 114, Appl

ALIGNMENTS

RESULT 1
US-09-887-593-2
; Sequence 2, Application US/09887593
; Patent No. US20020161212A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel B.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Leong, Kahan
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Saffran, Douglas C.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: BPC-1: A SECRETED BRAIN-SPECIFIC PROTEIN EXPRESSED AND
; TITLE OF INVENTION: SECRETED BY PROSTATE AND BLADDER CANCER CELLS
; FILE REFERENCE: 1703-017 US1
; CURRENT APPLICATION NUMBER: US/09/887,593
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 09/374,135
; PRIOR FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/095,982
; PRIOR FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-593-2

Query Match 100.0%; Score 862; DB 9; Length 158;
Best Local Similarity 100.0%; Pred. No. 44e-87;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIHRSVLIHVASLIILHLISGATKGTETSTSTKSVGCGTWTXKAEGLTSPNVS 60
Db 1 MIHRSVLIHVASLIILHLISGATKGTETSTSTKSVGCGTWTXKAEGLTSPNVS 60
QY 61 KYPRDRCIYITAEAPQCEIYFDEKYSLEPSGECKFHIEVVDGPRGSPITIGRPGQ 120

```
DB      61 KYPDPRECTIIEAAPROCTIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGRCQ 120
QY      121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTPGK 158
DB      121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTPGK 158

RESULT 2
US-10-138-588-38
; Sequence 38, Application US/10138588
; Publication No. US20040018594A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: NOVEL ANTIBODIES THAT BIND TO ANTIGENIC POLYPEPTIDES, NUCLEIC ACI
; FILE REFERENCE: 21402-347A
; CURRENT APPLICATION NUMBER: US/10/138,588
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/288,395
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/308,901
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/313,388
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/324,757
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/288,900
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 203
; SEQ ID NO 38
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-138-588-38

Query Match      98.7%; Score 851; DB 15; Length 526;
Best Local Similarity 100.0%; Pred. No. 3.4e-85;
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MHGSRVLIHVASLIILHSGATKKGTETKOTTSVOCGWTWTHAAGGIFTSNPNYS 60
DB      1 MHGSRVLIHVASLIILHSGATKKGTETKOTTSVOCGWTWTHAAGGIFTSNPNYS 60
QY      61 KYPDPRECTIIEAAPROCTIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGRCQ 120
DB      61 KYPDPRECTIIEAAPROCTIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGRCQ 120
QY      121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTP 156
DB      121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTP 156

RESULT 3
US-10-041-406-2
; Sequence 2, Application US/10041406
; Publication No. US20020164705A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; TITLE OF INVENTION: 39362, A NOVEL CUB DOMAIN CONTAINING
; FILE REFERENCE: 10448-130001
; CURRENT APPLICATION NUMBER: US/10/041,406
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: 60/260,286
; PRIOR FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-406-2
```

```
Query Match      98.7%; Score 851; DB 13; Length 533;
Best Local Similarity 100.0%; Pred. No. 3.4e-85;
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MHGSRVLIHVASLIILHSGATKKGTETKOTTSVOCGWTWTHAAGGIFTSNPNYS 60
DB      1 MHGSRVLIHVASLIILHSGATKKGTETKOTTSVOCGWTWTHAAGGIFTSNPNYS 60
QY      61 KYPDPRECTIIEAAPROCTIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGRCQ 120
DB      61 KYPDPRECTIIEAAPROCTIEFYDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGRCQ 120
QY      121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTP 156
DB      121 QNPVTKSSGRFLMIKFFADGELESMGFSARYNFTP 156

RESULT 4
US-10-162-435-27
; Sequence 27, Application US/10162435
; Publication No. US20030096305A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Leiby, Kevin R.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE-ASSOCIATED PROTEIN AND
; FILE REFERENCE: 10448-189001
; CURRENT APPLICATION NUMBER: US/10/162,435
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/836,499
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: PCT/US01/12420
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/197,507
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/891,008
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: PCT/US01/19963
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/214,220
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/860,868
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/16013
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,674
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/886,429
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: PCT/US01/20055
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/213,963
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 10/041,406
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: PCT/US02/00275
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/260,286
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: US 09/934,268
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/41811
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,612
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 533
```



```
;; PRIOR FILING DATE: 2000-06-28
;; PRIOR APPLICATION NUMBER: 60/217,487
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/225,758
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/220,963
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: 60/217,496
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/225,447
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/218,290
;; PRIOR FILING DATE: 2000-07-14
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 192
;; SOFTWARE: Patent Ver. 2.0
;; SEQ ID NO 116
;; LENGTH: 141
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-242-747-116
```

```
Query Match 55.6%; Score 479; DB 15; Length 141;
Best Local Similarity 66.1%; Pred. No. 8.4e-45;
Matches 80; Conservative 21; Mismatches 20; Indels 0; Gaps 0;
```

```
QY 38 SVCCGTAKAEGIFTSPPVPSKYPDRRCIYIIAAPQCIETLYFDEKYSIPESECK 97
DB 9 ATCCGIVWRNSNGHFASPPYPSDYPNKECIYIIAAPQRIELTFDEHYIEPSFECR 68
QY 98 FDHIEVDGPFSPFIIGRFGQGNPPVIRKSGRFLMIKEFADGEISMGFSAHYNFTPG 157
DB 69 FDHIEVDGPFSPFLIDRYCGVSPPLINSTGRFMWIKKSSDELEGLGFRAYKSTPLG 128
```

```
QY 158 K 158
DB 129 K 129
```

```
RESULT 9
US-10-276-774-2238
;; Sequence 2238, Application US/10276774
;; Publication No. US20040053245A1
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; TITLE OF INVENTION: Tom et al
;; FILE REFERENCE: 21272-030
;; CURRENT APPLICATION NUMBER: US/10/276,774
;; PRIOR FILING DATE: 2002-11-18
;; PRIOR APPLICATION NUMBER: 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; NUMBER OF SEQ ID NOS: 2700
;; SOFTWARE: Custom
;; SEQ ID NO 2238
;; LENGTH: 352
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-276-774-2238
```

```
Query Match 54.5%; Score 470; DB 12; Length 352;
Best Local Similarity 54.9%; Pred. No. 2.6e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
```

```
QY 6 SVLHI--VASLIILHLSGATKKGTSEKQTSFQSVCCGWTAKAEGIFTSPPVPSKYP 63
DB 28 SVLKVLLITVAVVGIAVAAQKTQDGNIGIKHLPATCCGIVWRNSNGHFASPPYPSYP 87
QY 64 PDRRCIYIIAAPQCIETLYFDEKYSIPESECKFDHIEVDGPFSPFIIGRFGQGNP 123
DB 88 PNKECIYIIAAPQRIELTFDEHYIEPSFECRFDHIEVDGPFSPFLIDRYCGVSP 147
```

```
QY 124 PVIKSGRFLMIKEFADGEISMGFSAHYNFTP 156
DB 148 PLIRSTGRFMWIKKSSDELEGLGFRAYKSTPL 180
```

```
RESULT 10
US-10-147-493-114
;; Sequence 114, Application US/10147493
;; Publication No. US20040029217A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Beresini, Maureen
;; APPLICANT: DeForge, Laura
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Sherwood, Steven
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P330R1G345
;; CURRENT APPLICATION NUMBER: US/10/147,493
;; PRIOR FILING DATE: 2002-05-17
;; Prior Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 114
;; LENGTH: 525
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-147-493-114
```

```
Query Match 54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4.4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
```

```
QY 6 SVLHI--VASLIILHLSGATKKGTSEKQTSFQSVCCGWTAKAEGIFTSPPVPSKYP 63
DB 8 SVLKVLLITVAVVGIAVAAQKTQDGNIGIKHLPATCCGIVWRNSNGHFASPPYPSYP 67
QY 64 PDRRCIYIIAAPQCIETLYFDEKYSIPESECKFDHIEVDGPFSPFIIGRFGQGNP 123
DB 68 PNKECIYIIAAPQRIELTFDEHYIEPSFECRFDHIEVDGPFSPFLIDRYCGVSP 127
QY 124 PVIKSGRFLMIKEFADGEISMGFSAHYNFTP 156
DB 128 PLIRSTGRFMWIKKSSDELEGLGFRAYKSTPL 160
```

```
RESULT 11
US-10-145-127-114
;; Sequence 114, Application US/10145127
;; Publication No. US20040033558A1
;; GENERAL INFORMATION:
```

```
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Beresini, Maureen
;; APPLICANT: DeForge, Laura
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
```



```

; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C252
; CURRENT APPLICATION NUMBER: US/10/145,127
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 114
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-145-127-114

```

```

Query Match          54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4,4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

```

```

QY 6 SVLHI--VASLIILHSATKKGTEKOTSETOKSVOCGWTWKAAEGIFTSPNPSKYP 63
DB 8 SVLKVLITLVVVEGIAVAQKTODGONIGIKHLPATQCGIWRKSNHGHPASPNYDPSY 67
QY 64 PDRECIYIIAARPROCIIEYFDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCGOQNP 123
DB 68 PNKECIYIIAARPROCIIEYFDEHYIEPSFEKCFDHEVRDGPFGSPFLIDRYCGVXSP 127
QY 124 PVIKSSGRFLMIKFFADGELSMGFSARVNFPT 156
DB 128 PLIRSTGRFMWIKFSSDELEGLGFRAXKYSFIP 160

```

```

RESULT 12
US-10-160-503-114
; Sequence 114, Application US/10160503
; Publication No. US20040033559A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C446
; CURRENT APPLICATION NUMBER: US/10/160,503
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 114
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-160-503-114

```

```

Query Match          54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4,4e-43;

```

```

Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;
QY 6 SVLHI--VASLIILHSATKKGTEKOTSETOKSVOCGWTWKAAEGIFTSPNPSKYP 63
DB 8 SVLKVLITLVVVEGIAVAQKTODGONIGIKHLPATQCGIWRKSNHGHPASPNYDPSY 67
QY 64 PDRECIYIIAARPROCIIEYFDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCGOQNP 123
DB 68 PNKECIYIIAARPROCIIEYFDEHYIEPSFEKCFDHEVRDGPFGSPFLIDRYCGVXSP 127
QY 124 PVIKSSGRFLMIKFFADGELSMGFSARVNFPT 156
DB 128 PLIRSTGRFMWIKFSSDELEGLGFRAXKYSFIP 160

```

```

RESULT 13
US-10-143-118-114
; Sequence 114, Application US/10143118
; Publication No. US20040038335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C228
; CURRENT APPLICATION NUMBER: US/10/143,118
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 114
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-143-118-114

```

```

Query Match          54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4,4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

```

```

QY 6 SVLHI--VASLIILHSATKKGTEKOTSETOKSVOCGWTWKAAEGIFTSPNPSKYP 63
DB 8 SVLKVLITLVVVEGIAVAQKTODGONIGIKHLPATQCGIWRKSNHGHPASPNYDPSY 67
QY 64 PDRECIYIIAARPROCIIEYFDEKYSIEPSWECKFDHIEVRDGPFGSPPIIGFCGOQNP 123
DB 68 PNKECIYIIAARPROCIIEYFDEHYIEPSFEKCFDHEVRDGPFGSPFLIDRYCGVXSP 127
QY 124 PVIKSSGRFLMIKFFADGELSMGFSARVNFPT 156
DB 128 PLIRSTGRFMWIKFSSDELEGLGFRAXKYSFIP 160

```

```

RESULT 14
US-10-144-993-114
; Sequence 114, Application US/10144993
; Publication No. US20040038336A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

```
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C261
CURRENT FILING DATE: 2002-05-13
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 114
LENGTH: 525
TYPE: PRT
ORGANISM: Homo Sapien
US-10-144-993-114

Query Match      54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4.4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

QY 6 SVLHI--VASLIIHLGATKGTGKQTSETQSVQCGTWTGHAEGIFTSPNYSKRP 63
DB 8 SVLKVLITVLYVEGLAVAGKQTQDGNIGKHIPATQCGIWTSTNGHFASPNYDSTP 67
QY 64 PDRECIYITFAARQCIETLYFDEKYSIEPSWECKFPHIEVRDGPFGSPPIIGFCCGQNP 123
DB 68 PNKECIYITFAARQRIETLTFDEHYIIEPSFECKFPHLEVRDGPFGSPPIIDRYCGVXSP 127
QY 124 PVTKSSGRFLWIKFPADGELSMGFSARYNFTP 156
DB 128 PLIRSTGRFWIKFSSDELEGLGFRKYSFIP 160

RESULT 15
US-10-158-787-114
; Sequence 114, Application US/10158787
; Publication No. US20040039164A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C449
CURRENT FILING DATE: 2003-04-03
```

```
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 114
LENGTH: 525
TYPE: PRT
ORGANISM: Homo Sapien
US-10-158-787-114

Query Match      54.5%; Score 470; DB 12; Length 525;
Best Local Similarity 54.9%; Pred. No. 4.4e-43;
Matches 84; Conservative 29; Mismatches 38; Indels 2; Gaps 1;

QY 6 SVLHI--VASLIIHLGATKGTGKQTSETQSVQCGTWTGHAEGIFTSPNYSKRP 63
DB 8 SVLKVLITVLYVEGLAVAGKQTQDGNIGKHIPATQCGIWTSTNGHFASPNYDSTP 67
QY 64 PDRECIYITFAARQCIETLYFDEKYSIEPSWECKFPHIEVRDGPFGSPPIIGFCCGQNP 123
DB 68 PNKECIYITFAARQRIETLTFDEHYIIEPSFECKFPHLEVRDGPFGSPPIIDRYCGVXSP 127
QY 124 PVTKSSGRFLWIKFPADGELSMGFSARYNFTP 156
DB 128 PLIRSTGRFWIKFSSDELEGLGFRKYSFIP 160

Search completed: May 18, 2004, 18:29:35
Job time : 44 secs
```